

Strategie di formazione delle parole in varietà iniziali di polacco L2*

Jacopo Saturno

1. Introduzione

Chiunque abbia studiato una lingua straniera si sarà certo trovato nella situazione di non conoscere una specifica parola, indispensabile per farsi capire in un dato contesto. Sin dalla scuola vengono spesso insegnate tecniche per evitare o almeno risolvere questi difficili episodi, quali ad esempio ricorrere a sinonimi o parafrasi, oppure tipicamente ‘riformulare la frase’. Ci sono però casi in cui la competenza nella lingua straniera non è ancora sufficiente per mettere in atto tali strategie, oppure semplicemente la parola è troppo specifica e non è possibile indicarne il significato neanche con una certa approssimazione.

In luogo di porci nella prospettiva didattica, in cui l’obiettivo è la correttezza della produzione linguistica in termini di corrispondenza alla lingua bersaglio, affrontiamo l’argomento dalla prospettiva della linguistica acquisizionale di stampo funzionalista (Croft 2001), in cui la produzione di chi parla è guidata dall’obiettivo comunicativo. In questo senso un enunciato è efficace se veicola in modo soddisfacente il significato voluto, indipendentemente dalla sua correttezza grammaticale.

Nell’ambito di tale approccio introduciamo la nozione di ‘varietà di apprendimento’, o ‘interlingua’, con la quale denotiamo tutti gli stadi che attraversa nel tempo la capacità di un individuo di utilizzare la lingua bersaglio (Perdue 1993). Analogamente, per ‘apprendente’ intendiamo una persona che per qualunque motivo si trovi ad imparare una nuova lingua in età adulta. Ciascuna varietà di apprendimento è vista come un sistema coerente governato da una grammatica, intesa come un insieme di regole grazie alle quali è possibile esprimere significati anche complessi con mezzi linguistici più o meno ridotti.

È piuttosto intuitivo che quanto più è embrionale la varietà di apprendimento, tanto più sarà sentito il problema delle lacune lessicali (De Angelis 2007: 33).

* Questo lavoro è stato finanziato nell’ambito del PRIN 2009 *Seconda lingua straniera nell’Europa multilingue: acquisizione, interazione, insegnamento*, coordinato da Giuliano Bernini. In particolare, l’autore è stato titolare di un assegno di ricerca dal titolo *Acquisizione di lingue seconde in classi italofone in condizioni di input controllato: per una prospettiva interlinguistica*. Per brevità, consideriamo una L2 qualunque lingua imparata in età adulta, indipendentemente dal numero di lingue conosciute e dal contesto in cui sono state acquisite. L1 indica naturalmente la lingua madre di un individuo.

Nelle prime fasi dell'acquisizione il lessico è necessariamente limitato, al che il sistema si adatta assegnando alle parole un significato più generico che non nelle varietà native, di modo che al medesimo elemento lessicale possono corrispondere più referenti leggermente diversi a seconda del contesto, per quanto tutti ovviamente accomunati da qualche tratto semantico. Può capitare tuttavia che sia richiesta una specifica parola senza la quale sarebbe difficile proseguire la comunicazione.

Concretizziamo il problema considerando uno dei test del progetto VIL-LA¹ (vedi oltre): a un gruppo di apprendenti iniziali di polacco L2 è chiesto di ripetere il contenuto di un breve filmato (*The Finite Story*, Dimroth 2012), il quale racconta di un incendio scoppiato nottetempo in una casa e di come grazie all'intervento dei pompieri tutto si risolve per il meglio. Per descrivere la scena in cui i pompieri spengono l'incendio, l'ideale sarebbe naturalmente conoscere la traduzione polacca della parola *spegnere*. Sappiamo però per certo che questo è impossibile per i nostri apprendenti: per risolvere tale difficoltà, diverse strategie possono allora essere adottate. Gli esempi da (1) a (4) riportano varie soluzioni adottate per descrivere la scena in cui i pompieri estinguono le fiamme con gli idranti. In primo luogo l'apprendente può semplicemente rinunciare alla comunicazione, rivolgendosi direttamente in italiano all'intervistatore (1):

(1) non so.

Come si è detto, tuttavia, le varietà di apprendimento dispongono sin dai livelli più iniziali di strategie utili a garantire l'efficacia della comunicazione anche in presenza di mezzi linguistici estremamente ridotti, come nel nostro caso. L'apprendente può innanzitutto scegliere di sfruttare il contesto, in misura in generale tanto più massiccia quanto più è lessicalmente povera l'interlingua. Oltre alla possibilità di riferirsi a un oggetto in maniera puramente deittica, cioè indicandolo fisicamente, se possibile, il contesto permette di utilizzare elementi lessicali noti e generici riferendoli a oggetti o situazioni inerenti alla situazione comunicativa condivisa con l'interlocutore (2).

- (2) a. *strażak jest sukces.*
pompieri è successo
- b. *i pożar jest koniec.*
e incendio è fine

¹ Desidero ringraziare sinceramente i due revisori anonimi per i loro suggerimenti e spunti, insieme a tutti i coordinatori e i membri del progetto VIL-LA; tra questi, in particolare, Marzena Watorek e Christine Dimroth hanno elaborato e coordinato i test *Question & Answer* e *Finite Story*, rispettivamente, mentre Giuliano Bernini, Ada Valentini, Roberta Grassi, Marina Chini e Cecilia Andorno hanno condotto le due edizioni italiane del progetto VIL-LA. Un ringraziamento particolare anche a tutti i partecipanti all'esperimento il cui entusiasmo e la cui pazienza hanno reso possibile la buona riuscita del progetto.

- c. *strażakiem pcha woda.*
pompieri spinge acqua
- d. *i pożar nie jest na dachu.*
e incendio non è su tetto
- e. *strażak pracuje.*
pompieri lavora

Negli esempi riportati in (2), gli enunciati sono facilmente comprensibili se l'interlocutore sa che si sta parlando di un incendio e che sono stati chiamati i pompieri per spegnerlo, tutti elementi introdotti precedentemente nel corso della narrazione.

Se però si esclude la possibilità di riferirsi al contesto, grazie al quale di fatto l'apprendente evita di dover ricorrere alla parola mancante, e si chiede di produrre un enunciato efficace anche in isolamento, allora rimane solo la possibilità di tentare di raggiungere l'obiettivo "inventando" una parola che abbia qualche probabilità di essere compresa dall'interlocutore. Naturalmente questo processo creativo non sarà arbitrario, ma piuttosto guidato da criteri di verosimiglianza e probabilità statistica, per quanto computati a un livello inconsapevole. Trovandosi nella necessità di integrare la propria competenza nella lingua bersaglio, l'apprendente ragionevolmente farà ricorso alle lingue che conosce, cioè la propria lingua madre (3a)-(3b) e altre lingue straniere (3c-d). Avremo quindi un caso di interferenza interlinguistica intenzionale, nello specifico un prestito *ad hoc* volto a risolvere una particolare situazione comunicativa.

- (3) a. *tragedia jest jest finito.*
tragedia è è *finito* (it)
- b. *strażak spegne pożar.*
pompieri *spegne* (it) incendio
- c. *pan czerwony is safe.*
signor Rossi *is safe* (ingl)
- d. *pożar nie [eksistiren].*
incendio non [eksistiren] (ted)

Possiamo però distinguere due tipi di prestito: in un caso la parola straniera è utilizzata nella sua forma originale, come negli esempi appena mostrati. In altre occasioni, più interessanti, è possibile cogliere dei processi di adattamento morfologico alla struttura tipica delle parole nella lingua bersaglio, o almeno all'idea che ne ha l'apprendente (4).

- (4) *eliminacja pożar.*
eliminazione incendio

In effetti nell'input compaiono numerose parole terminanti in *-acja*, ad esempio *kontynuacja* "continuazione", *prezentacja*, "presentazione" e altre, particolarmente salienti nel contesto del corso di lingua. L'enunciato in (4) propone un interessante esempio di tale strategia, particolarmente riuscita nel caso specifico: possiamo infatti affermare con certezza che l'apprendente non conosceva la parola *eliminacja* (vedi di seguito), la quale però esiste in polacco, con il significato, per quanto non rilevante, di 'selezione' (da un elenco o nell'ambito di un concorso). Il tentativo quindi si è rivelato fortunato, confermando l'esistenza di una logica sottostante a questa strategia. Possiamo infatti supporre che l'obiettivo dell'apprendente, scegliendo di adattare la morfologia del prestito alla lingua bersaglio, fosse di "indovinare" effettivamente una reale parola polacca, come nell'esempio appena discusso. Dal momento che la radice lessicale è data, in quanto semplicemente tratta dalla lingua madre², lo sforzo dell'apprendente si concentra sulla strategia di adattamento morfologico, cioè sulla scelta dei tratti segmentali (fonemi permessi nella lingua bersaglio), sovra-segmentali (es. posizione dell'accento), e degli eventuali formanti considerati più probabili nella L2.

Perché ciò avvenga, l'apprendente deve aver acquisito una conoscenza almeno parziale della morfologia della lingua bersaglio, in modo da potervi adattare la propria produzione. A questo scopo è necessario che analizzi l'input a cui è esposto in modo da estrarne le informazioni rilevanti, tra cui eventuali regolarità morfologiche in termini di flessione e formazione delle parole: tali regolarità a loro volta sono date dalla compresenza in una parola di una data forma (per esempio, un morfema derivazionale come *-ik* in *stolik*, 'tavolino') e un certo significato ('diminutivo'). La proposta di questo lavoro è che tale processo avvenga secondo un criterio probabilistico: Ellis (2006: 1) descrive gli apprendenti come "intuitive statisticians, weighing the likelihoods of interpretations and predicting which constructions are likely in the current context", e il processo di acquisizione come "the gathering of information about the relative frequencies of form-function mappings". Proveremo a verificare questa posizione quantificando la probabilità di associare una data forma alla funzione corrispondente, sfruttando le condizioni ideali offerte dall'input controllato del progetto VILLA.

2. Elaborazione dell'input in varietà di apprendimento iniziali

Uno studio quantitativo delle associazioni tra forma e funzione nell'input tuttavia pone notevoli difficoltà tecniche, tutte connesse all'estrema variabilità del contesto in cui ciascun apprendente impara la nuova lingua. Possiamo identificare due punti principali:

² Nel caso di sinonimi l'apprendente potrebbe anche interrogarsi su quale dei due termini abbia più probabilità di essere imparentato con la parola bersaglio.

- a) in primo luogo è piuttosto arduo riunire un numero sufficiente di partecipanti che partano esattamente dal medesimo livello di competenza nella lingua bersaglio; a livello teorico, inoltre, non è semplice definire in modo univoco che cosa sia tale livello di partenza e come sia possibile misurarlo.
- b) in condizioni naturalistiche, l'input a cui è esposto ciascun apprendente varia notevolmente per quantità e qualità, rendendo di fatto impossibile correlare in modo rigoroso questo fondamentale fattore con lo sviluppo della varietà di apprendimento. Per questo motivo di solito gli studi longitudinali hanno un approccio qualitativo più che quantitativo (Giacalone Ramat 2003; Perdue 1993). Altri studi hanno optato per la creazione di lingue artificiali *ad hoc* (Hulstijn 1997), la cui validità ecologica non è però del tutto chiara.

Il tema dell'elaborazione nell'input nelle varietà di apprendimento iniziali tuttavia è sempre più sentito nella ricerca acquisizionale, proprio perché stadi tanto embrionali di una L2 permettono di analizzare i meccanismi universali di acquisizione eliminando il filtro della competenza accumulata dall'apprendente nel suo percorso di apprendimento, e in una certa misura anche delle specificità della lingua madre e della lingua bersaglio (Rast 2008). Lo studio di questi stadi precoci insomma permetterebbe di accedere in maniera più diretta ai meccanismi universali dell'acquisizione, e in generale della facoltà di linguaggio umana.

Il progetto VILLA (*Varieties of Initial Learners in Language Acquisition*; Dimroth *et al.* 2013) si pone l'obiettivo di studiare queste fasi iniziali in maniera sperimentale, isolando quanto più possibile i processi di elaborazione autonoma dell'input da ogni tipo di interferenza con eventuali conoscenze linguistiche pregresse e altre variabili difficili da controllare.

Per far sì che tutti i partecipanti al progetto partissero dal medesimo livello di competenza nella L2, sono stati selezionati solo quei candidati che non avessero mai avuto esperienza né di polacco né di altre lingue slave, a causa delle evidenti somiglianze a livello grammaticale e lessicale, e laddove possibile nemmeno di altre lingue altamente flessive e morfologicamente complesse come greco, latino e tedesco. Oltre a evitare ogni competenza diretta nella lingua bersaglio, questi criteri miravano a evitare che i partecipanti avessero familiarità anche soltanto con i concetti metalinguistici di 'caso', 'paradigma' o 'aspetto', fra gli altri, che avrebbero potuto influenzare i naturali processi di elaborazione dell'input.

In secondo luogo era necessario assicurarsi che ciascun partecipante fosse esposto a input identico in qualità e quantità. A questo scopo i partecipanti hanno preso parte a un corso di polacco di 14 ore tenuto da un'insegnante madrelingua per 1 ora e 30 di lezione al giorno. Dal momento che i partecipanti non avevano alcuna esperienza di polacco, l'unico input possibile era costituito dal parlato dell'insegnante durante le ore di lezione e dalle eventuali presentazioni a video. Ai partecipanti era richiesto di non prendere appunti e di non cercare informazioni sulla lingua polacca al di fuori del corso, al fine di evitare ogni interferenza relativamente alla quantità o qualità dell'input a cui ciascuno era esposto. Solo a queste condizioni infatti è possibile isolare questa variabile così

sfuggente per concentrarsi sui processi comuni. Eventuali strategie condivise sarebbero dunque dovute alla struttura della facoltà di linguaggio umana e alle spontanee modalità di elaborazione dell'input.

La significatività del progetto VILLA quindi è di garantire il pieno controllo tipico degli studi sulle lingue artificiali mantenendo però le caratteristiche di una lingua storico-naturale in tutta la sua complessità. Nel caso del polacco la complessità morfologica del sistema nominale è evidente anche in un modello semplificato come il corpus VILLA: per quanto alcune classi flessive non compaiano che sporadicamente o siano del tutto assenti, è comunque possibile osservare il comportamento di apprendenti iniziali in relazione a caratteristiche tipiche delle lingue slave, quali il diffuso sincretismo tra terminazioni di caso o l'interazione tra genere di un sostantivo, terminazione del nominativo e paradigma flessivo.

L'esperimento è stato condotto in due edizioni in ciascuno dei cinque paesi coinvolti (Germania, Paesi Bassi, Francia, Regno Unito e Italia). Tutte le edizioni del corso si sono svolte seguendo la medesima metodologia e il medesimo programma, così da garantire la massima compatibilità dei dati ai fini del confronto interlinguistico. In questo capitolo ci concentriamo sui dati raccolti nelle due edizioni italiane, per un totale di 31 partecipanti.

3. *Il test Question & Answer*

Alla lezione seguivano diversi test di natura psicologica e linguistica, tra cui quello che verrà discusso in questo capitolo. Il test *Question & Answer* (*Q&A* d'ora in poi) è progettato per elicitarle le due diverse strutture copulari del polacco, *to jest* + NOM, 'questo è + NOM' e *on/ona jest* + STRUM, 'lui/lei è + STRUM', abbondantemente trattate durante il corso VILLA. Al partecipante è richiesto di descrivere un personaggio includendo in una struttura copulare le informazioni fornite. Dapprima viene presentata a schermo un'icona indicante il genere del referente; è importante precisare da subito che tutti gli elementi bersaglio del test si riferiscono a referenti umani, per cui la categoria grammaticale del genere coincide sempre con quella semantica del sesso biologico. Subito dopo all'apprendente è posta una domanda sotto forma di una delle tre strutture copulari possibili, cioè *kto to jest?* 'chi è questo³?', *kim on jest?* 'chi è lui?' oppure *kim ona jest?* 'chi è lei?'; infine viene presentata un'immagine rappresentante la nazionalità o professione del personaggio (FIGURA 1). Il compito del partecipante a questo punto è di esprimere quest'ultima informazione cercando al contempo di utilizzare la medesima struttura copulare utilizzata nella domanda, per esempio *on jest strażakiem* 'lui è un pompiere'.

³ La traduzione è necessariamente approssimativa, in quanto il dimostrativo *to* è invariabile rispetto al genere e al numero e come tale non ha un diretto equivalente in italiano.



Figura 1. Il test Question&Answer

Le risposte degli apprendenti sono state registrate in digitale e trascritte in IPA (cf. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*, Cambridge 1999) utilizzando i programmi ELAN (Brugman, Russell 2004) e CHAT/CLAN (MacWhinney 2000). Il test è stato proposto ai partecipanti in due occasioni, precisamente dopo 4:30 ore (T1) e dopo 10 ore (T2) di esposizione all'input. In questo capitolo ci limiteremo alla prima scadenza, particolarmente interessante a causa del periodo di acquisizione eccezionalmente ridotto.

Ci concentriamo inoltre sulla morfologia del solo nome prodotto dagli apprendenti, rimandando per una trattazione delle strutture morfosintattiche ad un'altra sede (Saturno 2015).

Possiamo identificare due dimensioni in base alle quali si differenziano gli elementi bersaglio del test *Q&A*: la prima è la trasparenza lessicale del nome, un parametro binario indicante se una data parola polacca è intuitivamente traducibile nella lingua madre degli apprendenti oppure no. Ad esempio *aktor* è immediatamente traducibile come 'attore' ed è quindi considerata una parola trasparente, laddove la parola *gasić* 'spegnere' è del tutto opaca per chi già non la conosca.

L'altro parametro è la frequenza dell'elemento bersaglio nell'input. Ai nostri fini è di nuovo sufficiente una classificazione binaria, per cui ci limitiamo a distinguere parole presenti e assenti nell'input. Ad esempio la parola *eliminacja* 'selezione' già citata è assente dall'input, ed è dunque indubbio che si tratti di un caso di elaborazione autonoma dall'apprendente. Il parametro delle frequenze è quello più rilevante per il nostro studio, in quanto se l'elemento lessicale non è presente nell'input, l'apprendente deve necessariamente 'inventarlo' facendo appello a tutte le proprie risorse linguistiche. Incrociando i due parametri si osserva inoltre che gli elementi bersaglio assenti sono anche trasparenti: ciò significa che se l'apprendente applica una strategia di formazione delle parole alla radice di una parola nota, per esempio tratta dalla propria lingua madre, avrà una

Elemento bersaglio	Genere	Pronuncia (IPA)	Traduzione
<i>australijczyk</i>	M	/aʊstra'lijtʃik/	'australiano'
<i>belgijka</i>	F	/bel'gijka/	'belga'
<i>kanadyjka</i>	F	/kana'dijka/	'canadese'
<i>pilot</i>	M	/'pilot/	'pilota'
<i>policjantka</i>	F	/poli'tsjantka/	'poliziotta'
<i>sekretarka</i>	F	/sekre'tarka/	'segretaria'

Tabella 1. Elementi bersaglio

ragionevole probabilità di produrre una parola polacca comprensibile. In ogni caso ciò che più ci interessa non è tanto il risultato, quanto piuttosto il processo, cioè il modo in cui, dato il materiale linguistico e la competenza a disposizione, l'apprendente cerca di ovviare a una lacuna lessicale.

Vogliamo inoltre eliminare ogni possibilità di interferenza tra elementi lessicali dell'input. A questo scopo un elemento bersaglio, pur classificato in sé come assente, è escluso dall'analisi se nell'input compare una parola imparentata. Il sostantivo *hiszpan*, 'spagnolo', per esempio, è escluso in quanto nell'input compare il corrispondente aggettivo *hiszpański*. Al termine del processo di selezione il corpus si limita a sei elementi bersaglio, per i quali è possibile escludere qualunque interferenza da parte di altri elementi lessicali (TABELLA 1):

4. Risultati

Considerando i sei elementi bersaglio per 31 apprendenti otteniamo un corpus atteso di 186 occorrenze totali. Escludendo le 32 omissioni, cioè le risposte in cui non viene prodotto alcun nome, il corpus si riduce a 154 enunciati, in cui di volta in volta possiamo riconoscere tre principali strategie:

- a) In 33 occasioni, l'apprendente produce un elemento lessicale diverso da quello richiesto ma compreso nell'input polacco. Si tratta normalmente dell'elemento semanticamente più prossimo, ad esempio *strażak* 'pompieri', per *policjantka* 'poliziotta' in ['ona 'jest stra'zakon]. Qui il tratto semantico comune è probabilmente rappresentato dal fatto che entrambi i referenti indossino un'uniforme e ricoprano una funzione pubblica.
- b) In 57 enunciati, l'apprendente produce una parola basata sulla radice di una lingua nota, spesso completata da una terminazione flessiva riconoscibile, ma senza utilizzare particolari formativi, ad esempio [sekretarjom] per *sekretarka* 'segretaria' in ['to 'jest sekre'tarjom].

- c) Infine, in 64 enunciati l'apprendente produce una parola basata su una radice nota, corredata da una terminazione flessiva riconoscibile e da un suffisso in *-k-*, ad esempio [vi'dzilkon] per *policjantka* 'poliziotto', in [ɔna'jestvidzi'lantko].

La TABELLA 2 mostra la distribuzione di queste soluzioni in base all'elemento bersaglio del test⁴.

I 64 casi in cui è riscontrabile l'uso del suffisso *-k-* corrispondono al 41% delle occorrenze totali e al 52% di tutte le occorrenze in cui non viene selezionato un altro elemento bersaglio tratto dall'input (scenari b e c). È chiaro, quindi, che nell'ambito delle genuine strategie di formazione della parole messe in atto dagli apprendenti, il suffisso *-k-* rappresenta di gran lunga la scelta più praticata. Solo 5 partecipanti su 31 (16%) non hanno prodotto alcuna occorrenza di *-k-* nelle loro risposte; per quanto riguarda i restanti, *-k-* compare da un minimo di una volta a un massimo di quattro sul totale delle sei risposte attese, con un valore di media e mediana pari a 2 (TABELLA 3).

Analizzando più nel dettaglio le risposte degli apprendenti (TABELLA 4) possiamo di nuovo identificare due principali dimensioni che descrivono la variabilità della produzione, in maniera paragonabile per tutti gli elementi bersaglio⁵.

In primo luogo può variare la radice lessicale selezionata dai partecipanti attingendo al repertorio della lingua madre o di altre lingue conosciute. Per esempio, per la parola bersaglio *policjantka* 'poliziotto', troviamo tipi ispirati alle radici /polis/, /polits/, /polif/, di origine presumibilmente italiana o inglese (cf. it. *polizia*, ingl. *police*), e altri formati su /vi'dzil/, più chiaramente ispirati all'italiano (cf. it. *vigile*).

La scelta della radice come si è detto non influenza la probabilità che la parola sia costruita con il formante *-k-*, e non è dunque di particolare interesse per il nostro studio. Nel quadro generale delle strategie a disposizione dell'apprendente per colmare le lacune lessicali, tuttavia, possiamo ipotizzare che alcune radici siano sentite come più probabili per ragioni psico-tipologiche, cioè della distanza percepita tra la propria lingua madre e la lingua bersaglio (Kellerman 1979). Se una parola è sentita come troppo comune (es. 'pilota'), l'apprendente potrebbe decidere di utilizzare un altro termine meno diffuso nell'intento di suonare meno 'italiano' (es. 'aviatore'; in effetti cf. pol. *awiator*). Dal momento che è piuttosto difficile trovare apprendenti perfettamente monolingui, cioè senza alcuna competenza in altre lingue, per quanto ridotta, il risultato sarà facilmente influenzato dalle relazioni reciproche di tutte le lingue note (Bardel, Lindqvist 2007); nel nostro esempio, quindi, una radice potrà essere favorita perché presente (oppure, viceversa, perché assente) in una

⁴ Per l'analisi statistica è stata usata la suite R (cf. *R: A Language and Environment for Statistical Computing*, <<http://www.R-project.org/>>; ultimo accesso: 12.11.2015).

⁵ Nella TABELLA 4, chiamiamo 'tipo' una forma a cui possono essere ricondotte, con solo minime variazioni fonologiche, una o più risposte degli apprendenti.

Elemento bersaglio	altro lemma (a)	non -k- (b)	-k- (c)	omissioni
<i>Australijczyk</i>	0	14	16	1
<i>Belgijka</i>	17	2	9	3
<i>Kanadyjka</i>	0	11	14	6
<i>Pilot</i>	2	15	6	8
<i>Policjantka</i>	5	7	11	8
<i>Sekretarka</i>	9	8	8	6
Totale	33	57	64	32

Tabella 2. Strategie di formazione delle parole per elemento bersaglio

occorrenze di -k-	0	1	2	3	4	5	6
numero di apprendenti	5	7	5	9	5	0	0

Tabella 3. Occorrenze di -k- per apprendente

Bersaglio	Tipi lessicali prodotti (IPA)
<i>Pilotem</i>	[pi'lotkom; militi'karom; a'vjetkou]
<i>Policjantka</i>	[po'liska; po'liskom; po'liskja; poli'tselkom; po'litfka; vi'dzilkon; vidzi'lanka; vidzi'lantko]
<i>Sekretarka</i>	[segre'tarka; segre'tarkjon; sekre'tarikom; sekre'terkon; infor'matika]
<i>Kanadyjka</i>	[kana'deskom; ka'nadkom; kana'dikon; kana'djenko; kana'djeka; ka'nadko; ka'nadikom; kana'deska; a'mirkou; kana'denka]
<i>Belgijka</i>	[bel'zike; 'belkom; 'belka; bel'dzika; 'bel'gikom; ni'tetska; 'bel'gika]
<i>Australijczyk</i>	[austra'ljankox; austra'likon; austra'lika; ays'traliko; ays'trinko; ays'traliko; os'tralka; austral'tjekon]

Tabella 4. Tipi lessicali prodotti dagli apprendenti

delle lingue conosciute dall'apprendente e a seconda delle sue relazioni con le altre presenti nel repertorio.

A questo proposito è interessante notare che nel valutare la trasparenza del lessico, e quindi la sua vicinanza alla lingua madre, gli apprendenti si basano non soltanto sulla forma della radice intesa come sequenza di fonemi, ma anche su tratti sovrasegmentali quali la posizione dell'accento. Sempre nell'ambito del progetto VILLA, Valentini e Grassi (in stampa) hanno mostrato come in un test di traduzione dal polacco i punteggi di trasparenza fossero molto più elevati per quelle forme di parola in cui l'accento cade sulla stessa sillaba dell'equivalente italiano (es. *fotograf*, /fo'tograf/, 'fotografo-NOM') che non nel caso contrario (es. *fotografem*, /foto'grafem/, 'fotografo-STRUM').

Tornando alle risposte degli apprendenti, una seconda dimensione della variazione riguarda la forma di parola: i tipi prodotti presentano solo due tipi di terminazione flessiva, le quali corrispondono alle due forme di parola richieste dal test, cioè *-a* /a/ NOM e *-q* /oŵ/ STRUM, naturalmente tenuto conto della variabilità fonetica caratteristica delle varietà di apprendimento iniziali.

Notiamo anche una tendenza del suffisso *-k-* con la relativa terminazione di genere femminile a sovraestendersi a referenti maschili, sul cui genere e sesso l'immagine non lascia dubbi. Relativamente ai personaggi della FIGURA 2, ad esempio, si riscontrano fra le tante le forme [pi'lɔtkom] o [a'vjetkou] per l'immagine a sinistra, e [austra:'likon] o [austral'fjekon] per quella a destra.



Figura 2. Referenti bersaglio di genere maschile nel test Q&A

5. *Discussione*

Sulla base dei dati appena presentati si può argomentare che il suffisso *-k-* sia stato identificato dagli apprendenti come un elemento caratteristico di alcune parole polacche, e che per questo venga incluso con evidente sistematicità anche nelle nuove forme prodotte durante il test. Si può quindi sostenere che sia stata elaborata una associazione tra questa forma e una particolare funzione, per quanto non necessariamente in accordo con le regole della lingua bersaglio.

Possiamo aprire la discussione dei risultati partendo da due fatti notevoli. In primo luogo, gli apprendenti si sono rivelati in grado di estrarre un formante dal flusso del parlato e usarlo in modo produttivo dopo sole 4 ore e 30 di esposizione all'input. In secondo luogo, è anche curioso che proprio *-k-* sia l'unico formante utilizzato dagli apprendenti, il che spinge a interrogarsi sul perché, fra quelli presenti nell'input, proprio questo suffisso sia stato identificato e non altri.

È bene ricordare qui che la varietà di polacco proposta durante il corso VILLA si discosta in parte dalle varietà native, o piuttosto ne costituisce un sottoinsieme in cui, fra le molte possibilità a disposizione, ne vengono realizzate solo alcune (Bernini 2015). In particolare, il lessico utilizzato nel corso è stato accuratamente pianificato e manipolato a scopi sperimentali, così che si può sostenere che le proporzioni degli elementi lessicali che vi compaiono non sempre rispecchiano quelle delle varietà native.

Il suffisso *-k-* è particolarmente diffuso e produttivo nelle varietà native di polacco, in cui concorre a formare anche altri formanti derivazionali dal significato più specifico, fra cui *-aczka*, *(ow)niczka*, *-owiczka*, *-i/ycielka*, *-arka*, *-anka*, *-ówka*, *-istka*. Nella sua forma più semplice il suffisso serve principalmente a formare *nomina actionis* (5a), *nomina patientis/objecti* (5b), e soprattutto diminutivi propri (5c) e etimologici (5d). La funzione più rilevante per il nostro lavoro è però quella di formativo di nomi umani femminili derivati da nomi maschili⁶ (5e) (Grzegorzczkowska 1979). Fra le funzioni elencate solo le ultime due mantengono a tutt'oggi un grado significativo di produttività (Jadacka 2005).

- (5) a. *walczyć*, “combattere” > *wal-ka* “lotta”
 b. *skakać*, “saltare” > *skakan-ka* “corda per saltare”
 c. *głowa*, “testa” > *głów-ka*, “testolina”
 d. *ręka*, “mano” > *rącz-ka*, “maniglia”
 e. *aktor*, “attore” > *aktor-ka* “attrice”.

Ci concentriamo sulla derivazione di nomi femminili dai corrispondenti maschili in quanto è l'unico significato rappresentato nella varietà semplificata di polacco dell'input VILLA. Ai nostri fini, tuttavia, mentre è indubbio che i nomi in *-k-* corrispondano a referenti umani femminili, non è sempre altrettanto

⁶ Il fatto che *-k-* con i suoi allomorfi sia di fatto l'unico suffisso derivazionale in grado di creare nomi femminili dai corrispondenti maschili ha qualche conseguenza sociolinguistica relativamente all'adeguamento del lessico alle mutate condizioni sociologiche. Fino all'inizio del XX secolo il sistema creava regolarmente nomi femminili a partire da nomi maschili utilizzando i suffissi derivazionali a disposizione. A partire dalla fine della prima guerra mondiale, però, e in maniera più marcata alla fine della seconda, i neologismi via via creati, e divenuti necessari in seguito all'accesso di molte donne a professioni fino ad allora tipicamente ‘maschili’, sono stati sentiti come inaccettabili o pragmaticamente marcati in senso derisorio o dispregiativo, il che ha portato alla necessità, sentita tutt'oggi, di creare nuovi nomi utilizzando strategie alternative (Jadacka 2005: 126-129).

evidente il sottostante processo di derivazione, dal momento che nella maggior parte dei casi mancano dall'input i corrispondenti primitivi di genere maschile. Per esempio, è assente dall'input VILLA il primitivo maschile *kucharz*, /'kuxaz/ 'cuoco', da cui si forma il derivato femminile *kucharka*, /ku'xarka/ 'cuoca'. Anche quando il primitivo maschile fosse presente, inoltre, spesso intervengono trasformazioni fonologiche a rendere meno evidente il processo di suffissazione, per esempio nella coppia *włoch*, /'vwox/ 'italiano' > *włoszka*, /'vwoʃka/ 'italiana'.

Tralasciando dunque la natura derivazionale del suffisso, è nostro obiettivo cercare di quantificare la probabilità che un apprendente possa associare alla forma *-k-* il significato di "nome di persona". A questo scopo, come operazione preliminare è necessario percorrere una breve rassegna di tutti i tipi di parola che possono includere *-k-*, inteso adesso non più come un morfema, ma soltanto come suono. Oltre ai nomi femminili di cui primariamente ci occupiamo, troviamo allora in primo luogo un ristretto numero di nomi maschili con nominativo singolare in *-k*, i quali assumono la terminazione *-a* al caso genitivo e talvolta anche accusativo, a seconda della classe nominale di appartenenza (6):

- | | | | | | |
|-----|----|--------------|---------------|----------------|-------------------|
| (6) | a. | <i>polak</i> | 'polacco-NOM' | <i>polak-a</i> | 'polacco-GEN/ACC' |
| | b. | <i>język</i> | 'lingua-NOM' | <i>język-a</i> | 'lingua-GEN' |

In *-(s)ka* terminano regolarmente gli aggettivi femminili il cui nominativo maschile è in *-(s)ki* (7):

- | | | | |
|-----|----|---------------------|-----------------|
| (7) | a. | <i>amerykańsk-a</i> | 'americana-NOM' |
| | b. | <i>wysok-a</i> | 'alta-NOM' |

Troviamo infine *-ka* come terminazione di III persona singolare di alcuni verbi molto frequenti (8):

- | | | |
|-----|----------------|---------|
| (8) | <i>mieszka</i> | 'abita' |
|-----|----------------|---------|

Mentre sono particolarmente rilevanti i casi in cui *-ka* denota nomi umani e in particolare femminili, per l'analisi quantitativa delle pagine seguenti dovremo tenere conto anche degli altri casi, a prima vista irrilevanti.

L'analisi che segue è ispirata agli studi condotti nell'ambito del paradigma teorico della Morfologia Naturale (Dressler 1987), in cui la 'naturalità' di una determinata struttura linguistica è una misura strettamente legata alla nozione di marcatezza, con cui si trova in una relazione di proporzionalità inversa. In generale, tra gli effetti di un alto grado di naturalità possiamo elencare: a) maggiore diffusione tra le lingue del mondo b) maggiore resistenza al cambiamento linguistico; c) maggiore facilità di elaborazione da parte di pazienti affetti da disturbi del linguaggio; d) maggiore facilità di elaborazione in tempo reale; e) precedenza nell'acquisizione, la caratteristica più direttamente rilevante per il nostro lavoro.

Il grado di naturalezza di una struttura linguistica può essere stabilito in base a 11 parametri (Crocco Galeas 1998), i quali a loro volta derivano da universali linguistici dettati dalle caratteristiche e limitazioni della facoltà di linguaggio umana. Ad esempio, il parametro della diagrammaticità prevede che una struttura sia tanto più naturale quanto più le sue parti (es. morfemi) sono chiaramente segmentabili e associabili a un significato univoco. Altri parametri invece vanno considerati sullo sfondo del particolare sistema linguistico per cui vengono calcolati (Wurzel 1987). Tra questi ultimi, due in particolare sono rilevanti per il nostro studio:

- a) uniformità: a una funzione corrisponde una e una sola forma;
- b) trasparenza: a una forma corrisponde una e una sola funzione.

Ciascuno caratterizza la relazione tra forma e funzione dal punto di vista dell'una o dell'altra. Considerando entrambi gli aspetti contemporaneamente otteniamo un altro parametro di natura più globale:

- c) biunivocità: forma e funzione si trovano in una relazione biunivoca (cioè ciascuna corrisponde univocamente all'altra).

Tutti i parametri della Morfologia Naturale sono scalari, cioè non pongono una scelta binaria tra 'naturale' e 'non naturale', ma piuttosto permettono di collocare una data struttura lungo un *continuum* di naturalezza. È stato suggerito che il grado di naturalezza in base a ciascuno degli 11 parametri sia misurabile, tanto da poter utilizzare la teoria in funzione sia esplicativa sia predittiva, per esempio relativamente alla frequenza di un determinato tratto tra le lingue del mondo (Dressler 1985).

Il paradigma della Morfologia Naturale è stato applicato con successo allo studio dell'acquisizione di L1, mostrando ad esempio che le prime produzioni linguistiche dei bambini sono particolarmente sensibili a parametri quali l'iconicità e la trasparenza (Dressler 2011). Affrontando il medesimo problema dell'associazione tra forma e funzione, la ricerca psicolinguistica ha elaborato modelli 'associazionisti' volti a spiegare i meccanismi di elaborazione del linguaggio in tempo reale, cioè come da segni linguistici sia possibile risalire a un significato e viceversa. Tra questi si segnala il *Competition Model* (MacWhinney, Bates 1989), il cui principale strumento metodologico è denominato *cue validity*: tale parametro misura statisticamente la probabilità che una data funzione venga associata alla forma corrispondente e viceversa. La prospettiva in cui si pone il modello è quella dell'elaborazione del parlato: forme e funzioni sono associate in base a un criterio probabilistico, per cui quanto più una forma è presente quando viene espresso un dato significato, tanto più forte sarà il legame fra i due. Si parla dunque di *task*, cioè situazioni in cui la funzione è presente, e di *clues*, 'indizi', cioè di strutture linguistiche che la segnalano. Ad esempio, nel caso della funzione di 'soggetto' in italiano, l'insieme delle *task* sarà dato da tutti gli enunciati in cui è presente un soggetto, e gli indizi, più o meno 'validi', saranno tutte le strut-

ture linguistiche che possono segnalare il costituente con questa funzione, fra cui l'accordo col verbo, la posizione iniziale nell'enunciato, ecc. I medesimi indizi possono essere più o meno validi a seconda della lingua considerata: nel caso dell'esempio precedente, in italiano l'indizio più efficace sembra essere l'accordo col verbo, mentre in inglese è la prima posizione nell'enunciato.

Il modello propone che la validità di un indizio sia misurabile, e in particolare che possa essere calcolata come il prodotto di due indici:

- a) diffusione (*Cue Availability*): indica la frequenza con cui la forma è presente in concomitanza con la funzione bersaglio. Questo parametro si può calcolare come il rapporto tra i casi in cui la forma è presente e segnala una funzione e i casi totali in cui è presente la funzione.
- b) affidabilità (*Cue Reliability*): indica l'affidabilità con cui una data forma segnala la funzione in esame. Il parametro è espresso dal rapporto tra i casi in cui la forma segnala correttamente la funzione e i casi totali in cui compare la forma.

Tale metodologia si è rivelata efficace nel descrivere le dinamiche di apprendimento di molte lingue, tra cui una lingua strutturalmente simile al polacco come il russo. In uno studio comparativo, Kempe e MacWhinney (1998) mostrano che nonostante la maggiore complessità della morfologia nominale del russo rispetto a quella del tedesco, il sistema dei casi del primo viene acquisito più rapidamente e con maggior precisione. Gli autori concludono che ciò sia una conseguenza del fatto che il sistema di paradigmi del russo, per quanto ricco e complesso, sia al contempo anche più univoco, cioè sia più facile associare a una data forma linguistica (es. terminazione di caso) un certo significato (es. soggetto). Nel caso del tedesco invece il sincretismo diffuso e la generale ambiguità delle terminazioni rendono l'acquisizione più lenta e difficile. In altre parole, la *Cue Validity* delle terminazioni di caso del russo è più elevata di quella del tedesco, portando così a una loro più rapida ed efficace acquisizione.

Suggeriamo qui che gli strumenti metodologici del *Competition Model* possano essere compatibili con il modello della Morfologia Naturale, permettendo anzi di quantificare in modo preciso la facilità di elaborazione di una particolare struttura. Il parallelo vuole che se in una lingua una struttura è più naturale come conseguenza di un elevato grado di biunivocità tra forma e funzione, allora anche la sua *Cue Validity* sarà più elevata, e viceversa.

In particolare, nell'ambito della Morfologia Naturale il parametro dell'uniformità segnala in che misura una data funzione è espressa da una e una sola forma. Ciò si può esprimere come il rapporto tra il numero di casi in cui la funzione è espressa da una data forma e quelli totali in cui compare la funzione, il che evidentemente nel *Competition Model* corrisponde alla misura della *Cue Availability*.

Analogamente, la trasparenza indica la misura in cui una data forma segnala una e una sola funzione: cioè può essere espresso in termini matematici come il rapporto tra il numero di casi in cui la forma segnala la funzione in oggetto e

la frequenza totale della forma. Come si vede, anche questa formulazione è sovrapponibile alla misura della *Cue Reliability*.

Tali misure tuttavia, seppur concettualmente semplici, nella pratica degli studi in condizioni naturalistiche sono difficilmente calcolabili a causa dell'impossibilità di quantificare tutte le occorrenze delle forme e funzioni a cui è esposto un apprendente durante il suo percorso di acquisizione. In molti studi questi valori sono calcolati sulla base di corpora scritti: Kempe e MacWhinney (1998) per esempio si affidano a una selezione di libri di testo. Si tratta però di approssimazioni, cioè di valori medi calcolati in un contesto didattico e sulla base della lingua scritta, per quanto talvolta strutturata a imitazione della varietà parlata.

La metodologia VILLA invece permette di verificare il modello in condizioni ideali. Come si è detto, infatti, a) tutti gli apprendenti sono stati esposti a input identico in termini di quantità e qualità, e b) l'input è stato interamente trascritto ed etichettato, rendendo così possibile quantificare in modo accurato la frequenza di forme, funzioni e eventuali sovrapposizioni.

La forma di nostro interesse è naturalmente il suffisso *-k-*, seguito dalla terminazione di caso (*-a* NOM oppure *-q* STRUM). Per quanto riguarda la funzione a cui la forma può essere associata, si può facilmente intuire che il significato in questione orbita intorno a quello di 'referente umano': ci si potrebbe chiedere tuttavia se tale significato non possa essere precisato ulteriormente avvicinandosi a quello del polacco, specificando ad esempio 'referente umano *femminile*'. Le risposte dei partecipanti (vedi Tabella 2) mostrano tuttavia che il suffisso viene utilizzato in maniera produttiva anche per descrivere i due elementi bersaglio di sesso chiaramente maschile (Figura 2). Tali occorrenze problematiche si potrebbero considerare come delle piccole deviazioni dall'associazione prevalente tra forma e funzione: potremmo ipotizzare che la possibilità di creare una parola dal suono autenticamente polacco sia prioritaria rispetto all'espressione accurata del genere, almeno nel contesto del test. È opportuno ricordare qui che l'unica fonte di informazioni sul polacco a disposizione degli apprendenti è costituita dall'input VILLA, nel quale, mancando con solo rare eccezioni i primitivi di genere maschile, la funzione derivazionale del suffisso *-k-* è del tutto oscurata. Di conseguenza si può escludere che i nomi femminili caratterizzati da questo suffisso siano sentiti come più marcati, in termini di naturalezza, rispetto ai loro equivalenti maschili, che dovrebbero essere del tutto sconosciuti ai partecipanti⁷. Inoltre, se da un lato è vero che *-k-* per ragioni di frequenza caratterizza principalmente i nomi di persona femminili, è altrettanto vero che lo stesso suono ricorre talvolta anche con nomi di persona maschili, come ad esempio *muzyk* 'musicista', o *chińczyk* 'cinese'.

Per maggior sicurezza in ogni caso discuteremo qui l'associazione di *-k-* con il significato più generale, cioè quello di 'nome di persona'. I valori relativi al significato di 'nome di persona *femminile*' sono comunque riportati nella TABELLA

⁷ Desidero ringraziare qui uno dei revisori anonimi per aver segnalato la necessità di questa importante precisazione.

nomi di persona	1339	nomi umani femminili	616
parole in <i>-k-</i>	1558	parole in <i>-k-</i>	1558
nomi umani in <i>-k-</i>	809	nomi femminili in <i>-k-</i>	610
<i>Cue Availability</i> /trasparenza	52%	<i>Cue Availability</i> /trasparenza	39%
<i>Cue Reliability</i> /uniformità	60%	<i>Cue Reliability</i> /uniformità	99%
<i>Cue Validity</i> /biunivocità	31%	<i>Cue Validity</i> /biunivocità	38%

Tabella 5. *Cue validity*/biunivocità del suffisso *-k-*

5. A questo proposito, è interessante notare che la quasi totalità dei nomi femminili è costruita con il suffisso *-k-*. La tabella mostra le frequenze per la forma e la funzione in questione, insieme a una quantificazione dei parametri discussi.

Relativamente alla *cue availability*/trasparenza, troviamo 1558 occorrenze di parole in *-k-* (inclusendo ogni classe di parola⁸), di cui 809 descrivono referenti umani, il che risulta in un rapporto del 52%. In termini statistici ciò significa che incontrando una parola in *-k-* si ha il 52% delle probabilità che si tratti di un nome umano: questo indicatore misura dunque l'univocità con cui la forma in esame rimanda al relativo significato. Passando alla *cue reliability*/uniformità, gli 809 nomi umani in *-k-* sono il 60% di tutti i nomi umani ($n = 1339$): di nuovo, ciò significa che dato un nome umano, nel 60% dei casi terminerà in *-k-*.

Moltiplicando i due indicatori si ottiene la misura globale della *Cue Validity*/biunivocità, che nel nostro caso è del 31%. Questo valore non ha un significato intrinseco, ma può essere paragonato ad altri calcolati secondo lo stesso procedimento per elementi linguistici simili, cioè altri suffissi derivazionali.

Per calcolare tali valori è necessario dapprima individuare eventuali altri formanti di nomi umani presenti nell'input VILLA. La TABELLA 6 raggruppa i sostantivi rilevanti in base al genere/sexo e alla forma del nominativo e dello strumentale singolare. Per ciascun lemma viene anche fornita la frequenza dopo 4 ore e 30 di esposizione all'input, cioè al momento del test. Dal momento che 8 è considerato il numero minimo di occorrenze necessarie affinché un elemento possa essere recepito e ricordato (Collins *et al.* 2009: 344 e riferimenti), sono riportati solo i sostantivi con frequenza uguale o superiore a questa soglia⁹.

Dal punto di vista della formazione delle parole, possiamo identificare alcune categorie in cui si dividono i nomi umani dell'input VILLA. In un primo gruppo troviamo i nomi maschili il cui nominativo singolare termina in conso-

⁸ È stata esclusa soltanto la profrase *tak*, 'si', in quanto in una lingua come il polacco il suo significato è tanto evidente quanto slegato dalla categoria del genere.

⁹ Tra i sostantivi esclusi, alcuni si trovano in un rapporto di primitivo-derivato con altri elementi di gran lunga più comuni, es. *kucharz* 'cuoco' (freq. 1) > *kuchar-ka* 'cuoca' (freq. 43). Tuttavia, la loro scarsissima diffusione (spesso un'unica occorrenza o poco più) sembra escludere la possibilità che gli apprendenti possano avere identificato il rapporto che lega i due sostantivi, e di conseguenza la funzione del suffisso *-k-*.

		M				F	
		-iec/-cem		-k/-kiem		-ka/-ką	
Lemma	Freq	Lemma	Freq	Lemma	Freq	Lemma	Freq
<i>Chorwat</i> 'croato'	61	<i>chłopiec</i> 'ragazzo'	8	<i>chińczyk</i> 'cinese'	53	<i>aktorka</i> 'attrice'	32
<i>Fotograf</i> 'fotografo'	35	<i>niemiec</i> 'tedesco'	72	<i>muzyk</i> 'musicista'	37	<i>artystka</i> 'artista (f)'	33
<i>Francuz</i> 'francese'	43	<i>sportowiec</i> 'sportivo'	9	<i>polak</i> 'polacco'	59	<i>brazylijka</i> 'braziliana'	43
<i>Inżynier</i> 'ingegnere'	39			<i>strażak</i> 'pompieri'	45	<i>dziewczynka</i> 'bambina'	9
<i>Lekarz</i> 'medico'	54					<i>kucharka</i> 'cuoca'	43
<i>Listonosz</i> 'postino'	53					<i>nauczycielka</i> 'insegnante (f)'	66
<i>Norweg</i> 'norvegese'	47					<i>niemka</i> 'tedesca'	58
<i>Profesor</i> 'professore'	39					<i>polka</i> 'polacca'	108
<i>Student</i> 'studente'	25					<i>portugalka</i> 'portoghese (f)'	42
<i>Tłumacz</i> 'traduttore'	10					<i>studentka</i> 'studentessa'	46
<i>Włoch</i> 'italiano'	20					<i>tłumaczka</i> 'traduttrice'	56
						<i>włoszka</i> 'italiana'	67

Tabella 6. Classi flessive nell'input VILLA

nomi virili	723	nomi virili	723
parole in <i>-ec/-cem</i>	95	parole in <i>-k/-kiem</i>	885
nomi virili in <i>-ec/-cem</i>	90	nomi virili in <i>-k/-kiem</i>	199
<i>Cue Availability</i> /trasparenza	12%	<i>Cue Availability</i> /trasparenza	22%
<i>Cue Reliability</i> /uniformità	95%	<i>Cue Reliability</i> /uniformità	27%
<i>Cue Validity</i> /biunivocità	12%	<i>Cue Validity</i> /biunivocità	6%

Tabella 7. *Cue Validity*/biunivocità del suffisso *-iec* e *-k*

nante (-C), o piuttosto con la terminazione zero. Ciò dà luogo a una discreta variabilità nella forma di questi nomi, che non necessariamente l'apprendente è in grado di raggruppare in una singola categoria in base alla terminazione: quello di 'consonante' è un concetto metalinguistico con cui non è detto che i partecipanti al progetto VILLA avessero dimestichezza.

Possiamo però identificare due sottogruppi caratterizzati da un comportamento più regolare grazie all'intervento di un morfema derivazionale (TABELLA 7). Il primo comprende nomi maschili in *-iec* come *niemiec*, /'njemjets/ 'tedesco'; il secondo nomi in *-k* come *strażak*, /'strazak/ 'pompieri'. In entrambi i casi è facile riunire tali nomi in un'unica classe a causa della terminazione comune; tuttavia a questi gruppi appartengono ben pochi tipi lessicali, con la conseguenza che il paradigma corrispondente è relativamente raro nell'input. Per questo motivo anche qualora la trasparenza/*Cue Reliability* delle terminazioni *-iec* o *-k* risultino elevate, comunque il valore globale della *Cue Validity* rimane piuttosto basso a causa della scarsa frequenza nell'input.

In conclusione, quindi, il fatto che proprio il suffisso *-k* sia stato selezionato come marca dei nomi umani non sorprende eccessivamente, se si considera la sua discreta frequenza nell'input e il buon livello di biunivocità tra la forma e il significato che veicola. Sorprende piuttosto che tale regolarità sia stata individuata da quasi tutti gli apprendenti e dopo sole 4 ore e 30 di esposizione all'input, senza alcun tipo di istruzione esplicita.

6. Conclusioni

In questo lavoro abbiamo cercato di mostrare come l'associazione tra forma e funzione possa essere facilitata dall'intervento di particolari morfemi derivazionali, i quali hanno l'effetto di rendere più riconoscibili come appartenenti a un'unica classe un gran numero di parole accomunate da un particolare tratto semantico. Nel caso in esame, i partecipanti all'esperimento VILLA si trovavano di fronte al compito di colmare le loro lacune lessicali utilizzando come materiale grezzo da una parte il limitatissimo input della lingua bersaglio, dall'altra

la loro competenza nelle lingue note, cioè la lingua madre ed eventuali lingue straniere. Gli apprendenti si sono dimostrati in grado di produrre nuove parole sfruttando il suffisso *-k-*, chiaramente estratto dall'input e utilizzato in modo produttivo e relativamente appropriato. In realtà solo il fatto che talvolta venga sovraesteso per caratterizzare anche referenti chiaramente maschili ci impedisce di affermare che l'associazione tra forma e funzione sia conforme a quella dell'input bersaglio, e cioè che nella maggior parte dei casi denoti nomi animati femminili. Bisogna tuttavia considerare anche che nessun formativo per nomi maschili era abbastanza saliente nell'input per essere rilevato dai partecipanti in un tempo così breve, o almeno non abbiamo modo di affermare il contrario. L'uso di *-k-* anche con referenti non appropriati dunque potrebbe essere un tentativo più o meno consapevole di approssimazione, motivato dalla priorità del significato referenziale rispetto a quello grammaticale.

Dal punto di vista dell'apprendente, l'uso di un formante estratto dall'input ha probabilmente lo scopo di adattare la forma di un elemento lessicale appartenente a un'altra lingua - la lingua madre o un'altra lingua straniera - al fine di renderlo più "polacco". Questo processo è del tutto indipendente dal ruolo che il suffisso *-k-* ricopre nelle varietà native, in quanto nell'input VILLA, mancando i primitivi maschili, la derivazione non è evidente. È bene però precisare che ci sono alcune eccezioni a questa tendenza. Nella tabella 6 troviamo almeno due coppie di sostantivi frequenti in cui è evidente il processo derivazionale: *student*, 'studente' > *studentka* 'studentessa', e *thumacz*, 'traduttore' > *thumaczka*, 'traduttrice'. In tre altre coppie la derivazione, pur presente, è meno evidente a causa di trasformazioni fonologiche: *wloch* /'vwox/, 'italiano' > *włoz-ka* /'vwoʃka/, 'italiana', *niemiec* /'n'em'ets/, 'tedesco' > *niem-ka* /'n'emka/, 'tedesca', e *polak* /'polak/, 'polacco' > *pol-ka* /'polka/, 'polacca'. Sulla base di questi dati si potrebbe pensare che il suffisso *-k-*, più che alla generica idea di 'nome di persona', sia stato associato alla sua funzione derivazionale, o almeno al significato più specifico di 'nome di persona *femminile*'. Mentre non è possibile escludere definitivamente questa possibilità, ci sono tuttavia alcuni argomenti contrari. In primo luogo, come già ricordato, il suffisso *-k-* viene utilizzato in modo sistematico anche per descrivere referenti maschili. In secondo luogo, il test richiedeva non tanto di derivare nuove parole a partire da un primitivo polacco dato, quanto piuttosto di creare nuovi elementi lessicali sulla base dell'input VILLA e di tutte le conoscenze linguistiche precedenti, ossia la L1 e eventuali lingue straniere. È vero che nel test Q&A le quattro parole bersaglio di genere femminile sono dei derivati sia in polacco, sia nella L1 degli apprendenti, e quindi, in termini di morfologia naturale, sono più marcate rispetto ai primitivi maschili; tuttavia ciò non sembra sufficiente per affermare che anche nella L2 gli apprendenti abbiano voluto creare un derivato, e non semplicemente una parola la quale, sulla base della loro analisi dell'input, avesse un suono il più possibile 'polacco'. A rigore, però, il test non pare adeguato per distinguere tra le due spiegazioni concorrenti che potrebbero motivare la scelta del suffisso *-k-* tra i tanti disponibili, e cioè da un lato il suo ruolo derivazionale e/o la forte associazione con i nomi femminili, dall'altro il rapporto che legherebbe questa

forma ai nomi umani in generale. Come che sia, ci interessa qui la forza dell'associazione tra forma e funzione, che può essere quantificata nel modo mostrato in questo lavoro: nei fatti, ciò che cambia tra le due spiegazioni è la funzione a cui si cerca di associare la forma, cioè 'nome animato femminile' in un caso, e genericamente 'nome animato' nell'altro.

Come che sia, abbiamo cercato di mostrare che la scelta di un particolare formante è funzione della sua biunivocità, nei termini della Morfologia Naturale, o della sua *Cue Validity*, secondo il *Competition Model*. Più un elemento è diffuso nell'input e più è trasparente riguardo alla funzione che codifica, più facilmente verrà notato dagli apprendenti e utilizzato in caso di necessità.

Un'ipotesi che si potrebbe proporre specularmente vuole che il suffisso *-k-* faciliterebbe anche l'acquisizione della morfologia flessiva mediante una più efficace assegnazione degli elementi lessicali al genere appropriato. Nell'input VILLA, infatti, tutti i nomi femminili presentano le terminazioni *-a* e *-q* al nominativo e allo strumentale rispettivamente, e appartengono al medesimo paradigma. Osservando i tipi di parole prodotti dai partecipanti, notiamo che le terminazioni prodotte sono nella maggior parte dei casi una corretta approssimazione di una legittima forma di parola: *-a* NOM oppure *-q* STRUM. La presenza di un suffisso particolarmente saliente permetterebbe cioè agli apprendenti di riunire le parole in cui compare in un'unica classe nominale e di individuarne le regolarità nella flessione, cioè appunto il paradigma. Dal punto di vista morfologico, infatti, nell'input nativo il suffisso è sempre associato a una terminazione pertinente. Un simile ruolo in effetti è stato suggerito per i diminutivi nel processo di acquisizione di varie lingue prime, tra cui lingue slave come russo (Kempe *et al.* 2007a), serbo (Ševa *et al.* 2007) e lo stesso polacco (Dąbrowska 2006). Lo stesso effetto è stato confermato anche nell'apprendimento di lingue straniere da parte di adulti (Kempe, Brooks 2008; Kempe *et al.* 2007). Una prima argomentazione vuole che i diminutivi contribuirebbero a rendere trasparenti rispetto al genere alcuni primitivi ambigui, come i nomi russi in consonante palatalizzata, i quali possono essere tanto maschili quanto femminili: vedi ad esempio *kost*, 'osso' FEM e *korabl'*, 'nave' MASC. Più ancora, i diminutivi ridurrebbero grandemente la variabilità della terminazione: laddove infatti i primitivi presentano una terminazione di caso con molte possibili realizzazioni fonetiche, la forma del diminutivo presenta un'unica terminazione facilmente identificabile. Così ad esempio in russo e polacco le molte e varie terminazioni del nominativo dei nomi maschili, cioè la maggior parte delle consonanti, possono essere ridotte ad una sola terminazione mediante l'uso del suffisso diminutivo *-ik*: tale regolarità favorirebbe poi l'associazione probabilistica del lemma con il genere corrispondente, in modo misurabile in termini di biunivocità come si è cercato di mostrare in questo lavoro. A causa del diffuso sincretismo della morfologia nominale slava, infine, la corretta disambiguazione del genere è a sua volta vitale anche per una corretta elaborazione delle terminazioni di caso, dal momento che solo conoscendo il genere di un dato nome è possibile disambiguare alcune terminazioni (Kempe, Brooks 2005): in polacco, ad esempio, la terminazione *-a* può indicare il nominativo femminile singolare (*gwiazda*, 'stella'), il genitivo

singolare maschile e neutro (*kota* ‘gatto’; *drzewa* ‘albero’) o il caso retto plurale dei nomi neutri (*krzesła* ‘sedie’).

Un ultimo risultato degno di nota è che i risultati qui presentati sono stati raggiunti in sole 4 ore e 30 di esposizione all’input, un tempo certamente brevissimo ma a quanto pare sufficiente affinché gli apprendenti riuscissero a segmentare l’input individuando radici lessicali e formanti. La trasparenza lessicale di alcune delle parole costruite col suffisso *-k-* può certamente essere stata di aiuto a questo scopo: nel caso di *artystka* ‘artista’, per esempio, è facile individuare una radice nota, da cui estrarre il significato lessicale generico; il partecipante può poi facilmente identificare l’elemento *-ka* e interrogarsi sul suo significato. Possiamo ipotizzare che ulteriore esposizione all’input permetterà infine di riconoscere il medesimo elemento anche in parole non trasparenti, come ad esempio *nauczycielka* ‘insegnante’.

In conclusione, i risultati di questo studio mostrano che anche dopo un tempo brevissimo di esposizione a un input del tutto sconosciuto, gli apprendenti si sono mostrati in grado di operare un’accurata analisi morfologica, riuscendo a replicare in modo produttivo un particolare tipo di struttura morfologica ricorrente in numerose parole polacche. Questo lavoro ha cercato di mostrare che tutto ciò è stato reso possibile, o almeno facilitato, dall’alto grado di biunivocità che l’uso di un suffisso derivazionale conferisce all’associazione tra una forma e la funzione corrispondente.

Bibliografia

- Bardel, Lindqvist 2007: C. Bardel, C. Lindqvist, *The Role of Proficiency and Psychotypology in Lexical Cross-Linguistic Influence. A Study of a Multilingual Learner of Italian L3*, in: M. Chini, P. Desideri, M. E. Favilla, G. Pallotti (a cura di), *Atti del VI Congresso Inter-nazionale dell’Associazione Italiana di Linguistica Applicata*, Perugia 2007, pp. 123-145.
- Bernini 2015: G. Bernini, *Il primo confronto con una norma: percezione e analisi dell’input iniziale in L2*, “Bulletin suisse de linguistique appliquée”, 2015, pp. 151-169.
- Brugman, Russell 2004: H. Brugman, A. Russell, *Annotating Multimedia/Multimodal Resources with ELAN*, in: M.T. Lino, M.F. Xavier, F. Ferreira, R. Costa, R. Silva (a cura di), *Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation*, Paris 2004, pp. 2065-2068.
- Collins et al. 2009: L. Collins, P. Trofimovich, J. White, W. Cardoso, M. Horst, *Some Input on the Easy/Difficult Grammar*

- Question: An Empirical Study*, "The Modern Language Journal", XCIII, 2009, 3, pp. 336-353.
- Crocco Galeas 1998: G. Crocco Galeas, *The Parameters of Natural Morphology*, Padova 1998.
- Croft 2001: W. Croft, *Grammar: Functional Approaches*, in: N.J. Baltes, P.B. Smelser (a cura di), *International Encyclopedia of the Social & Behavioral Sciences*, Oxford 2001, pp. 6323-6330.
- Dąbrowska 2006: E. Dąbrowska, *Low-Level Schemas or General Rules? The Role of Diminutives in the Acquisition of Polish Case Inflections*, "Language Sciences", XXVIII, 2006, 1, pp. 120-135.
- De Angelis 2007: G. De Angelis, *Third Or Additional Language Acquisition*, Clevedon 2007.
- Dimroth 2012: C. Dimroth, *Videoclips zur Elizitation von Erzählungen: Methodische Überlegungen und einige Ergebnisse am Beispiel der "Finite Story"*, in: B. Ahrenholz (a cura di), *Einblicke in die Zweitspracherwerbsforschung und ihre methodischen Verfahren*, Berlin 2012, pp. 77-98.
- Dimroth et al. 2013: C. Dimroth, R. Rast, M. Starren, M. Wątarek, *Methods for Studying the Learning of a New Language Under Controlled Input Conditions: The VILLA Project*, "EU-ROSLA Yearbook", XIII, 2013, pp. 109-138.
- Dressler 1985: W.U. Dressler, *On the predictiveness of Natural Morphology*, "Journal of Linguistics", XXI, 1985, 2, pp. 321-337.
- Dressler 1987: W.U. Dressler (a cura di), *Leitmotifs in Natural Morphology*, Amsterdam-Philadelphia 1987.
- Dressler 2011: W.U. Dressler, *The Rise of Complexity in Inflectional Morphology*, "Poznan Studies in Contemporary Linguistics", XLVII, 2011, 2, pp. 159-176.
- Ellis 2006: N. C. Ellis, *Language Acquisition as Rational Contingency Learning*, "Applied Linguistics", XXVII, 2006, 1, pp. 1-24.
- Giacalone Ramat 2003: A. Giacalone Ramat, *Verso l'italiano: percorsi e strategie di acquisizione*, Roma 2003.
- Grzegorzczkova 1979: R. Grzegorzczkova, *Zarys słowotwórstwa polskiego: słowotwórstwo opisowe*, Warszawa 1979.
- Hulstijn 1997: J. H. Hulstijn (a cura di), *Second Language Acquisition Research in the Laboratory*, Cambridge 1997 (= "Studies in Second Language Acquisition", XIX, 2)

- Jadacka 2005: H. Jadacka, *Kultura języka polskiego: fleksja, słowotwórstwo, składnia*, Warszawa 2005.
- Kellerman 1979: E. Kellerman, *Transfer and Non-Transfer: Where We Are Now*, "Studies in Second Language Acquisition", II, 1979, 1, pp. 37-57.
- Kempe, Brooks 2005: V. Kempe, P. J. Brooks, *The Role of Diminutives in the Acquisition of Russian Gender: Can Elements of Child-Directed Speech Aid in Learning Morphology?*, "Language Learning", LV, 2005, S1, pp. 139-176.
- Kempe, Brooks 2008: V. Kempe, P. J. Brooks, *Second Language Learning of Complex Inflectional Systems*, "Language Learning", LVIII, 2008, 4, pp. 703-746.
- Kempe et al. 2007: V. Kempe, P. J. Brooks, S. Gillis, G. Samson, *Diminutives Facilitate Word Segmentation in Natural Speech: Cross-Linguistic Evidence*, "Memory & Cognition", XXXV, 2007, 4, pp. 762-773.
- Kempe et al. 2007a: V. Kempe, P. J. Brooks, N. Mironova, A. Pershukova, O. Fedorova, *Playing with Word Endings: Morphological Variation in the Learning of Russian Noun Inflections*, "British Journal of Developmental Psychology", XXV, 2007, 1, pp. 55-77.
- Kempe, MacWhinney 1998: V. Kempe, B. MacWhinney, *The Acquisition of Case Marking by Adult Learners of Russian and German*, "Studies in Second Language Acquisition", XX, 1998, 3, pp. 543-587.
- MacWhinney 2000: B. MacWhinney, *The CHILDES Project: Transcription Format and Programs*, Mahwah 2000³.
- MacWhinney, Bates 1989: B. MacWhinney, E. Bates, *The Crosslinguistic Study of Sentence Processing*, Cambridge 1989.
- Perdue 1993: C. Perdue (a cura di), *Adult Language Acquisition, I. Field Methods: Cross-Linguistic Perspectives*, Cambridge 1993.
- Rast 2008: R. Rast, *Foreign Language Input: Initial Processing*, Clevedon 2008.
- Saturno 2015: J. Saturno, *Copular Structures in Polish L2*, "Linguistica e Filologia", XXXV, 2015, pp. 69-98.
- Ševa et al. 2007: N. Ševa, V. Kempe, P. J. Brooks, N. Mironova, A. Pershukova, O. Fedorova, *Crosslinguistic Evidence for the Diminutive Advantage: Gender Agreement in Russian and Serbian Children*, "Journal of Child Language", XXXIV, 2007, 1, pp. 111-131.

- Valentini, Grassi in stampa: A. Valentini, R. Grassi, *Oltre la frequenza. L'impatto della trasparenza e dell'accento sull'apprendimento del lessico in L2*, in: L. Corrà (a cura di), *Sviluppo della competenza lessicale. Acquisizione, apprendimento, insegnamento*, Roma in stampa.
- Wurzel 1987: W. U. Wurzel, *System-Dependent Morphological Naturalness in Inflection*, in: W. U. Dressler (a cura di), *Leitmotifs in Natural Morphology*, Amsterdam-Philadelphia 1987, pp. 59-98.

Abstract

Jacopo Saturno

Autonomous Word Formation Strategies in Initial Polish L2

This paper is devoted to a strategy of word formation as observed in the earliest stages of the acquisition of Polish L2. Within the VILLA project, 31 adult Italian L1 learners with no prior experience of Slavic languages took part in a 14-hour Polish course taught by a professional teacher, whose speech constitutes the only Polish input available to the learners. All input was planned, audio- and video- recorded, and finally transcribed and morphologically tagged: this way it is possible to thoroughly correlate the development of the interlanguage with the relevant features of the input that learners were exposed to.

The data show that after only 4:30 hours of exposure, the learners can produce new words using the derivational suffix *-k-*, which in the VILLA input is characteristic of the vast majority of feminine nouns (e.g. *kuchar-k-a*, "cook").

This indicates that after only minimal exposure to an unknown language, adult learners managed to identify the morphological structure of a class of nouns and apply it productively in their own output. Such results are explained in terms of input analysis, using the theoretical insights of Natural Morphology and the methodological tools of the Competition Model.