

Inferenza interlinguistica e intercomprensione nella Romània: sul ruolo della morfologia flessiva

1. Introduzione

In questo articolo presentiamo i risultati di una indagine empirica sulla inferenza interlinguistica ('interlingual inferencing' nel senso di Berthele 2011) tra romeno e italiano, ovvero sulla capacità da parte di parlanti nativi di lingua italiana o romena di riconoscere parole affini nella lingua "altra" senza avere conoscenza previa della lingua stessa (processo chiamato anche 'cognate guessing task' in Vanhove/Berthele 2015). Come si mostra nell'esempio (1), per parole affini ('cognates') intendiamo parole che abbiano lo stesso etimo latino, forma simile (pur se in gradi differenti) e uguale¹ significato in entrambe le lingue².

- | | | | |
|-----|-----------------|---|--|
| (1) | a. Lat. LUPUS | > | It. <i>lupo</i> / Rom. <i>lup</i> |
| | b. Lat. NUCES | > | It. <i>nocci</i> / Rom. <i>nuci</i> |
| | c. Lat. SAL | > | It. <i>sale</i> / Rom. <i>sare</i> |
| | d. Lat. CADITIS | > | It. <i>cadete</i> / Rom. <i>cădeți</i> |
| | e. Lat. MORIMUR | > | It. <i>moriamo</i> / Rom. <i>murim</i> |

Lo studio dell'inferenza interlinguistica si colloca a sua volta nel più ampio contesto di ricerca sull'intercomprensione e sul multilinguismo ricettivo, ovvero la situazione comunicativa nella quale «gli interlocutori si esprimono in lingue diverse pur comprendendosi, sia allo scritto sia all'orale, sia in presenza sia a distanza, sia in modalità sincrone sia in modalità asincrone» (Bonvino/Jamet 2016, 9)³. Infatti, l'abilità da parte di un parlante di percepire la somiglianza formale tra strutture linguistiche è un fattore fondamentale nell'intercomprensione e più in generale nei processi di acquisizione di L2 (si veda il concetto di 'transfer' in Ringbom 2006 per esempio).

Gli studi recenti che si sono occupati di inferenza interlinguistica (Berthele/Lambelet 2009, Berthele 2011, Vanhove/Berthele 2015, 2017) hanno proposto che la performance nel 'cognate guessing' dipenda da due ordini di fattori: (i) caratteristiche dei parlanti che si sottopongono al test come ad esempio la competenza multilinguistica e la capacità cognitiva di astrarre schemi

¹ Il problema della totale sinonimia nel processo di traduzione è ben noto (Newmark 1981, §8); ma considerata la natura dell'indagine (in cui una delle due lingue è sconosciuta), la questione non ha una diretta rilevanza nel nostro impianto teorico.

² Il nostro uso del termine 'cognate' come parola che discende da uno stesso etimo è quello generalmente usato negli studi di linguistica storica. Negli studi di psicolinguistica, invece, il termine non implica necessariamente una relazione su base etimologica (cf. Dijkstra/Van Heuven 2018, § 6.2.3).

³ Ad oggi, gli studi sull'intercomprensione si sono occupati di diversi aspetti, come ad esempio (i) il ruolo dell'intercomprensione nella didattica delle lingue romanze (Blanche-Benveniste/Valli 1997, Balboni 2005, 2009, Caddéo/Jamet 2013, Benucci 2015, Bonvino 2015); (ii) la misurazione dell'intelligibilità funzionale e/o percepita (Gooskens 2013); (iii) i fattori linguistici ed extralinguistici che favoriscono o ostacolano l'intelligibilità come ad esempio l'esposizione che il parlante ha avuto all' "altra" lingua e l'atteggiamento positivo o negativo nei confronti dell'"altra" lingua (cf. Schüppert 2011). Per quanto riguarda i fattori linguistici, ci si è concentrati su lessico (Van Bezooijen/Gooskens 2005), fonologia (Gooskens 2006, Kürschner *et al.* 2008) e ortografia (Van Bezooijen/Gooskens 2005, Kürschner *et al.* 2008). Il ruolo della morfologia (e della sintassi) non ha ancora ricevuto il dovuto interesse sperimentale (fanno eccezione, per le lingue germaniche, gli studi di Hilton *et al.* 2013 e Heeringa *et al.* 2014).

in modo fluido⁴, e (ii) caratteristiche delle parole affini usate come stimolo nell'esperimento⁵ e in particolare la distanza formale tra tali parole misurata attraverso il logaritmo di Levenshtein (Levenshtein 1966).

In questo articolo ci concentreremo sulle caratteristiche delle parole affini nel processo di riconoscimento piuttosto che sulle caratteristiche dei parlanti. Inoltre, per ragioni di spazio, riporteremo e analizzeremo nella sezione §3 solo i risultati collettivi degli esperimenti condotti in Italia (Montecatini).

Per quanto ci consta, il nostro è il primo studio a occuparsi di inferenza interlinguistica con particolare enfasi sul concetto di lessema, radice lessicale e suffisso flessivo. Le domande che ci poniamo sono in particolare:

1. L'identità della radice è parametro sufficiente per il riconoscimento di una parola affine (a livello di lessema ma anche a livello di significato lessicale di una base nella derivazione) in lingue affini?
2. La 'sostanza fonologica' della radice o della parola flessa (a livello di numero di sillabe e di presenza o meno di nessi consonantici) ha un ruolo nel processo di riconoscimento?
3. Esistono tipi di allomorfia più trasparenti di altri, ovvero qual è il confine formale (se ne esiste uno) per la intelligibilità a livello di parola?

2. Metodologia

Per investigare il ruolo della morfologia flessiva nel processo di inferenza interlinguistica tra italiano e romeno abbiamo formulato un esperimento che comprende un questionario⁶ e quattro test (nella modalità scritta ed esclusivamente su supporto cartaceo) ovvero:

- i. un test principale di traduzione di parole isolate (senza contesto);
- ii. un esercizio di comprensione testuale (vero/falso);
- iii. la traduzione di un testo di livello B2/B1 (utilizzato da Gooskens e collaboratori nel progetto Micrela);
- iv. un esercizio di tipo 'cloze test' sulla base di un testo per metà in italiano e per metà in romeno.

La batteria di test (incluso il questionario) della durata di un'ora complessivamente è stata somministrata in Italia a un gruppo di 113 studenti di madre lingua italiana (18-19 anni) del Liceo C. Salutati a Montecatini (Toscana), di cui 50,4% maschi e 49,6% femmine. In Romania i partecipanti erano 116 studenti di madre lingua romena (18 anni) del Liceu Teoretic Vasile Alecsandri a Iași (Moldavia), di cui 29,9% maschi e 70,1% femmine.

⁴ Le cosiddette 'participant-related variables' in Vanhove/Berthele 2015. Gli studi mostrano come il parlante multilingue con un ampio repertorio (con conoscenza di più lingue straniere) abbia una performance migliore negli esperimenti di 'cognate guessing'. Questo ha indotto gli studiosi ad affermare che la competenza multilingue conferisce una maggiore tolleranza per la distanza formale (Berthele 2011, Otwinowska-Kasztelanica 2011).

⁵ Gli 'item-related variables' che si sono dimostrati rilevanti statisticamente negli studi di Berthele e colleghi sono le differenze tra parole affini a livello di consonanti e soprattutto al principio di parola (fino alla prima consonante o 'onset' in Berthele 2011 e Möller/Zeevaert 2010). È interessante notare come nel nostro studio l'enfasi sul concetto di radice lessicale consenta di comprendere la ragione del ruolo prominente che riveste l' 'onset' di parola.

⁶ Il questionario mira ad acquisire informazioni sul background linguistico dei partecipanti e sul loro atteggiamento (positivo/negativo) nei confronti della propria e dell' "altra" lingua.

Il test (i) di traduzione di parole isolate rappresenta il nucleo stesso dell'esperimento e su di esso ci concentreremo in questo articolo. La decisione di sottoporre al parlante una lista di parole senza un contesto naturale è del tutto intenzionale nel 'cognate guessing task' (cf. Berthele 2011). Come afferma lo stesso Berthele infatti, l'assenza di un contesto naturale in questo caso, contrariamente a quanto accade generalmente nell'analisi linguistica, è infatti un vantaggio perché permette al ricercatore di investigare il solo ruolo della forma astraendo da fattori di natura semantico-pragmatica.⁷

La lista del test italiano comprende 166 parole romene in totale di cui 109 parole affini (stessa etimologia e stesso significato, es. *lup, mână, timp*) e 50 parole di origine non latina senza corrispondente in italiano, es. *abur, alifie, boarfă, iad, pagubă, pisică*. Mentre la lista del test somministrato in Romania comprende 183 parole italiane in totale di cui 109 parole affini e 73 parole italiane senza corrispondente affine in romeno, es. *barile, bosco, gatta, gambero*.

I criteri formali adottati nella selezione delle parole sono i seguenti: (a) il numero di sillabe della parola flessa (1, 2 o 3) per testare il ruolo della sostanza fonologica nel riconoscimento (maggiore la lunghezza della parola più facile il riconoscimento). (b) La presenza o meno di nessi consonantici nella radice lessicale per appurare se questi abbiano un effetto facilitatore nel riconoscimento (si veda la proposta di Vanhove/Berthele 2015 che le consonanti possano contribuire maggiormente al riconoscimento di parole affini rispetto alle vocali⁸). (c) La totale identità vs mancata identità della radice lessicale (per le parole affini) per testare l'idea che la totale identità della radice sia un parametro sufficiente nel processo di riconoscimento. (d) l'identità della desinenza (per le parole affini) vs assenza della desinenza (in Italia) vs mancata identità della desinenza. Si vedano gli esempi in (2) tratti dalla lista di parole romene somministrata in Italia.

(2)

- i. Gruppo A (= radice/ = suffisso): *capră, casă* (bisillabi con o senza nessi consonantici); *lacrimă, salută* (trisillabi con o senza nessi consonantici)
- ii. Gruppo B (= radice/ no suffisso): *corp, nas* (monosillabi con o senza nessi consonantici); *complic, dedic* (bisillabi con o senza nessi consonantici)
- iii. Gruppo C (= radice/ ≠ suffisso): *aspru, navă* (bisillabi con o senza nessi consonantici); *cărnuri, ramură* (trisillabi con o nessi consonantici)
- iv. Gruppo D (≠ radice/ = suffisso): *doarme, sună* (bisillabi con o senza nessi consonantici); *răspunde, perete* (trisillabi con o senza nessi consonantici)
- v. Gruppo E (≠ radice/ no suffisso): *sulf, foc* (monosillabi con o senza nessi consonantici); *cunosc, vecin* (bisillabi con o senza nessi consonantici)
- vi. Gruppo F (≠ radice/ ≠ suffisso): *brațe, oase* (bisillabi con o senza nessi consonantici); *creșteți, auziți* (trisillabi con o senza nessi consonantici)

⁷ «In naturalistic situations, there would be additional, contextual information available that contributes in significant ways to this decision, most notably the semantic fit of the item within the co- and context. The artificial and decontextualized nature of the word list data investigated here deliberately neglects these important constraints, in order to focus on the purely formal, linguistic side of the process» (Berthele 2011, 216).

⁸ «The implication is that consonants may contribute more to the subjective transparency of cognate relationships than do vowels, possibly due to their greater information value» (Vanhove/Berthele 2015, 99).

Oltre ai criteri formali sopra elencati che forniscono dati qualitativi, abbiamo calcolato la distanza ortografica (non normalizzata⁹) delle parole affini in romeno e italiano secondo l'algoritmo di Levenshtein che computa il numero di operazioni (inserzione, cancellazione o cambiamento) necessarie per trasformare una parola 'x' dell'italiano nella parola affine 'y' del romeno (cf. Tabella 1).

Ad esempio le parole del gruppo A, essendo identiche (i diacritici del romeno non vengono considerati e i parlanti sono stati istruiti esplicitamente ad ignorarne la presenza), corrispondono a un valore zero. Parole come *corp/corpo* e *aspru/aspro* hanno entrambe un valore '1' nonostante *corp* implichi la cancellazione della vocale finale dell'italiano *corpo* mentre *aspru* implichi il cambiamento della vocale finale (*o>u*). Il valore più alto è '5' per due sole parole *coasem/cuciamo* e *plăcem/piacciamo*.

Valore Levenshtein (non normalizzato)	Parole affini romeno/italiano
0	cade/cade, cântă/canta, capră/capra, carne/carne, casă/casa, contemplă/contempla, continui/continui, crede/crede, dormi/dormi, educă/educa, evită/evita, lacrimă/lacrima, salută/saluta, stradă/strada, vede/vede, verde/verde, voce/voce
1	aspru/aspro, cad/cado, cădeți/cadete, câmp/campo, cazi/cadi, coadă/coda, complic/complico, consacru/consacro, convinge/convince, corp/corpo, cred/credo, cresc/cresco, crud/crudo, dedic/dedico, dinte/dente, doarme/dorme, dorm/dormo, evit/evito, imit/imito, lapte/latte, lup/lupo, mâini/mani, mână/mano, metru/metro, nas/naso, navă/nave, nod/nodo, nuci/noci, pastor/pastore, perete/parete, piatră/pietra, poartă/porta, porc/porco, sare/sale, ședeți/sedete, sper/spero, sună/suona, vedeți/vedete, verificați/verificate, verzi/verdi, vezi/vedi
2	câmpuri/campi, canaluri/canali, cărnuri/carni, contăm/contiamo, cozi/code, creșteți/crescete, cunosc/conosco, foc/fuoco, fruct/frutto, iarbă/erba, nasuri/nasi, oase/ossa, om/uomo, os/osso, pierd/perdo, popor/popolo, prezintă/presenta, răspunde/risponde, șed/siedo, studiază/studia, tauri/tori, timp/tempi, vând/vendo, vecin/vicino, venim/veniamo
3	auziți/udite, credem/crediamo, ghidez/guido, impun/impongo, lăpturi/latti, marmură/marmo, moară/muore, mor/muoio, murim/moriamo, oameni/uomini, operează/opera, pot/posso, ramură/ramo, săruri/sali, sulf/zolfo, taur/toro, timpuri/tempi, vaci/vacche
4	braț/braccio, brațe/braccia, ierburi/erbe, patru/quattro, plac/piaccio, poate/può
5	coasem/cuciamo, plăcem/piacciamo

Tabella 1. Distanza di Levenshtein delle parole affini

Nonostante il fatto che la distanza di Levenshtein fornisca dati puramente quantitativi, gli studi di Berthele e colleghi hanno osservato come essa rappresenti un efficace parametro diagnostico per il 'cognate guessing'.

⁹ Si vedano Heeringa 2004 e Schepens *et al.* 2012.

Sul fronte quantitativo, dati importanti vengono anche dalla frequenza nei corpora e dalla ‘basicità’ delle parole affini incluse nella lista. Purtroppo, a causa delle costrizioni derivate dalla scelta dei criteri formali sopra discussi, accanto a parole italiane con un alto tasso di frequenza (CoLFIS, Corpus e Lessico di Frequenza dell’Italiano Scritto) e basicità (NVdB, Nuovo Vocabolario di Base della lingua italiana) come *tempo*, *mano*, *uomo* e *casa* troviamo parole come *zolfo*, *imitare*, *canale* che hanno al contrario tassi di frequenza e basicità piuttosto bassi. Per quanto non sia stato possibile usare frequenza e basicità nella selezione delle parole, queste variabili sono comunque considerate nell’analisi dei dati in §3.

3. Presentazione e discussione dei dati

I dati della performance collettiva (misurata in % di traduzioni corrette) in Italia sono riportati nelle Tabelle 2 e 3. È importante sottolineare come nell’analisi abbiamo considerato corretta ogni traduzione che veicolasse il significato contenuto nella radice lessicale, quindi sia forme flesse dello stesso lessema (come per esempio *imita*, *imiti*, *imitare* per il romeno ‘imit’) sia forme derivate che condividano la stessa radice lessicale e semantica (come ad esempio *imitazione* per ‘imit’, o *annusare*, *nasale* per ‘nasuri’). Questa scelta riflette la nostra preoccupazione teorica circa la possibilità che l’identità della radice in una lingua affine sia un parametro sufficiente per il riconoscimento di una ‘parola’ affine (a livello di lessema ma anche a livello di significato lessicale di una base nella derivazione).

Come si può vedere dalla Tabella 2, nella maggior parte dei casi, l’identità della radice è sufficiente per il riconoscimento del significato lessicale. Nel caso di identità dell’intera forma flessa il riconoscimento è sempre superiore al 90% raggiungendo spesso il 100%.

Tuttavia, i nostri dati mostrano come la morfologia flessiva possa impedire l’intelligibilità quando rende difficile la corretta segmentazione morfologica tra radice e desinenza. Questo accade soprattutto in casi in cui la desinenza è bisillabica a livello ortografico¹⁰, ad es. *câmp-uri*, *canal-uri*, *căd-eți*, *nas-uri*. In questi casi possiamo affermare che sia la sostanza fonologica della desinenza piuttosto che quella della radice ad avere un ruolo determinante nella capacità del parlante di riconoscere il significato lessicale della parola flessa.

Per quanto riguarda i valori sulla distanza di Levenshtein, possiamo notare come a un valore basso (pari a zero o a ‘1’¹¹) sembra riflettere in modo adeguato un livello di performance alto nel ‘cognate guessing’. Questo accade sia nel caso vi sia un’operazione di cancellazione (it. *corpo* > rom. *corp*) sia in caso di cambiamento della vocale finale (it. *aspro* > rom. *aspru*).

	$\geq 50\%$		< 50%	
	$\geq 70\%$	69-50%	49-31%	$\leq 30\%$
= radice / = suffisso	capră (100%) carne (100%) casă (100%) salută (100%) verde (100%) voce (100%) cântă (98%)			

¹⁰ Al contrario, quando la parola flessa è pronunciata in isolamento la desinenza *-uri* con /r/ palatalizzata è monosillabica.

¹¹ Nelle Tabelle 2 e 3 la distanza di Levenshtein è indicata con il numero in alto a destra_[n], con l’eccezione del valore zero che non viene rappresentato.

	contemplă (98%) dormi (98%) stradă (98%) crede (96%) cade (94%) continui (94%) educă (94%) lacrimă (94%) evită (92%) vede (92%)			
= radice / no suffisso	corp ₁ (100%) dorm ₁ (100%) pastor ₁ (98%) crud ₁ (96%) câmp ₁ (94%) evit ₁ (94%) nas ₁ (93%) dedic ₁ (89%) cred ₁ (88%) cresc ₁ (88%) nod ₁ (83%) complic ₁ (74%) imit ₁ (76%) porc ₁ (74%)	cad ₁ (64%) sper ₁ (64%)	lup ₁ (45%)	
= radice / ≠ suffisso	verifica ₁ (98%) metru ₁ (96%) șede ₁ (93%) studiază (93%) aspru ₁ (88%) navă ₁ (85%) vede ₁ (85%) consacru ₁ (81%) credem ₃ (81%) operează (81%) mână ₁ (76%)	venim ₂ (69%) contăm ₂ (59%) căde ₁ (55%)	canaluri ₃ (42%) câmpuri ₂ (31%) cărniuri ₂ (31%) marmură ₃ (31%)	ramură ₃ (12%) nasuri ₂ (25%)

Tabella 2. Dati raccolti in Italia per le parole con identità della radice lessicale

Il ruolo fondamentale dell'identità della radice è dimostrato anche dalla bassa performance nel riconoscimento delle parole con radice differente (cf. Tabella 3). È interessante notare come la distanza di Levenshtein in questi casi sia predittiva (ovvero sia un parametro efficace per predire la performance di riconoscimento) solo quando è superiore al valore di '3'.

	≥ 50%		< 50%	
	≥ 70%	69-50%	49-31%	≤ 30%
≠ radice / = suffisso	răspunde ₂ (90%) convinge ₁ (87%) doarme ₁ (88%) nuci ₁ (85%) coadă ₁ (72%) poartă ₁ (70%)	perete ₁ (61%), tauri ₂ (54%)	dinte ₁ (45%) mâini ₁ (44%)	vezi ₁ (20%) piatră ₁ (18%) iarbă ₂ (16%) lapte ₁ (13%) oameni ₃ (12%) prezintă ₂ (8%)

				sare ₁ (3%) cazi ₁ (3%) sună ₁ (3%) verzi ₁ (0%)
≠ radice / no suffisso	fruct ₂ (92%) cunosc ₂ (90%) vecin ₂ (72%)	taur ₃ (54%) pierd ₂ (52%)	murim ₃ (44%) șed ₂ (38%) mor ₃ (33%) popor ₂ (32%)	foc ₂ (28%) om ₂ (24%) impun ₃ (16%) braț ₄ (14%) pot ₃ (11%) sulf ₃ (11%) os ₂ (10%) timp ₂ (7%) vând ₂ (5%) plac ₄ (3%)
≠ radice / ≠ suffisso				ierburi ₄ (20%) plăcem ₅ (20%) auzi ₃ (7%) brațe ₄ (7%) timpuri ₃ (7%) crește ₂ (5%) moară ₃ (5%) oase ₂ (5%) ghidez ₃ (3%) poate ₄ (3%) săruri ₃ (1%) vacii ₃ (1%) coasem ₅ (0%) cozi ₂ (0%) lăpturi ₃ (0%) patru ₄ (0%)

Tabella 3. Dati raccolti in Italia per le parole senza identità della radice lessicale

Un valore inferiore a ‘3’ al contrario non è di per sé parametro che rifletta in modo efficace i risultati del test e solo un’osservazione qualitativa può fare luce sui dati ottenuti. Per esempio, parole come *verzi/verdi*, *cazi/cadi*, *vezi/vedi* hanno un valore di Levenshtein pari a ‘1’ ma hanno una percentuale di riconoscimento di 0%, 3% e 20% rispettivamente mentre parole come *nuci/noci*, *coada/coda* e *poarta/porta* con lo stesso valore di Levenshtein pari a ‘1’ hanno delle percentuali di successo tra l’85% e il 70%.

È interessante notare come differenze a livello consonantico (piuttosto che la presenza di nessi consonantici) rappresentino un ostacolo maggiore al riconoscimento rispetto a differenze a livello vocalico. Lo si vede nel cambiamento /d/ > /z/ (ortograficamente «d» > «z») in *verzi/verdi*, *cazi/cadi*, *vezi/vedi* dovuto in romeno ad affricazione storica ma anche nel cambiamento /l/ > /r/ (ortograficamente «l» > «r») nelle parole *sare/sale* e *popor/popolo*. Un ulteriore esempio viene dalla differenza nella performance tra due forme verbali, entrambe trisillabiche con e senza cambiamento consonantico, come *răspunde/risponde* (90%) e *prezintă/presenta* (8%).

¹² Si noti come la lettera «z» corrisponda in italiano a /dz/ e non a /z/.

Questi risultati sono in linea con i risultati ottenuti da Berthele (2011) che mostrano come il ruolo delle consonanti¹³ nel ‘cognate guessing’ sia prominente rispetto a quello delle vocali. Un apparente contro-esempio a questa tendenza è rappresentato da *timp/tempo* che ha un livello di riconoscimento molto basso (7%) nonostante la corrispondenza consonantica ‘t-mp’. Osservando i risultati più attentamente tuttavia notiamo che il 61% dei parlanti associa il romeno *timp* alla parola italiana *timpano*. Possiamo ipotizzare quindi che nel momento stesso in cui il parlante identifica una parola nella propria L1 che presenta totale identità nella sillaba iniziale, specialmente se non esistono competitori (altre parole che iniziano per ‘timp-’ come nel caso dell’italiano *timpano*) la ricerca è conclusa, pur non essendoci identità della radice.

Ci sono altri casi in cui la corrispondenza consonantica non è di per sé sufficiente a garantire una buona performance. Si vedano ad esempio *piatră/pietra* e *sună/suona* (18% e 3%) o *iarbă/erba* (16%). Questo sembra indicare che, a livello vocalico, alcune differenze ostacolano il riconoscimento più di altre. In particolare, se paragoniamo *piatră/pietra* (18%) con *coadă/coda* (72%) o *poarta/porta* (70%) sembra evidente che in casi in cui sia necessario il solo cancellamento di una vocale in un dittongo, il tasso di riconoscimento è molto alto. L’apparente contro-esempio della parola *ierburi/erbe* si può motivare pensando all’effetto che abbiamo notato per il suffisso romeno ‘-uri’ nelle parole con identità di radice.

4. Conclusioni

In questo articolo abbiamo riportato parte dei risultati di un più ampio studio empirico sull’inferenza interlinguistica tra italiano e romeno.

Abbiamo sottolineato come il nostro sia il primo studio sull’inferenza interlinguistica che fondi il proprio impianto teorico su concetti della morfologia flessiva come ‘lessema’, ‘radice lessicale’ e ‘desinenza’.

Abbiamo proposto che la totale identità della radice lessicale sia sufficiente (ma non necessaria) per una buona performance nel ‘cognate guessing task’ purché la sostanza fonologica della desinenza flessiva non impedisca la corretta segmentazione tra radice e desinenza stessa. Questo accade ad esempio con le parole *câmpuri*, *canaluri*, *cădeți*, *nasuri* in cui la desinenza romena è bisillabica a livello ortografico.

Crediamo che il concetto di radice lessicale possa anche fare luce sui risultati ottenuti negli studi precedenti che sottolineano l’importanza delle corrispondenze ortografiche a inizio di parola (Berthele 2011, Möller/Zeevaert 2010). Nelle lingue analizzate in tali studi come nel nostro infatti, la parola è strutturata secondo lo schema ‘radice+desinenza’ e la radice è spesso di natura monosillabica.

Inoltre, i dati sui casi in cui non vi sia identità della radice lessicale confermano il ruolo prominente della corrispondenza consonantica rispetto a quella vocalica come già proposto da studi precedenti (Berthele 2011 in particolare) ma mostrano come vi sia una gerarchia all’interno della categoria ‘corrispondenze vocaliche’ tra processi differenti (cambiamento di vocale, cancellazione di vocali all’interno di dittongo vs aggiunta di vocale) e come questo aspetto meriti un approfondimento.

¹³ « This suggests an even role of consonants: if they are the same bad inferers have a relatively high chance of inferring the words correctly, but differences in vowels do not seem to trigger worse performances» (Berthele 2011, 208).

Infine, abbiamo osservato come un parametro puramente quantitativo come la distanza di Levenshtein possa predire il successo o l'insuccesso nel 'cognate guessing' quando il valore sia minimo (zero, 1) o massimo (4, 5). Tuttavia, negli altri casi il parametro numerico si dimostra insufficiente al fine di comprendere più a fondo quali aspetti formali della parola affini nella lingua sconosciuta favoriscono o impediscono l'intelligibilità.

Università di Oxford
Research Centre for Romance Linguistics
Chiara CAPPELLARO
Martin MAIDEN

Riferimenti bibliografici

- Balboni, Paolo, 2005. «L'intercomprensione tra le lingue romanze: un problema di politica linguistica», in: Benucci, Antonella, *Le lingue romanze: una guida per l'intercomprensione*, Torino, UTET, 3-14.
- Balboni, Paolo, 2009. «Per una glottodidattica dell'intercomunicazione romanza», in: Jamet, Marie-Christine, *Orale e intercomprensione tra lingue romanze, ricerche e implicazioni didattiche*, Venezia, Cafoscarina, 197-203.
- Benucci, Antonella, 2015. *L'intercomprensione: il contributo italiano* Torino, UTET.
- Berthele, Raphael, 2011. «On abduction in receptive multilingualism», *Applied Linguistics Review* 2, 191-220.
- Berthele, Raphael / Lambelet, Amelia, 2009. «Approche empirique de l'intercompréhension: répertoires, processus et résultats», *Lidil. Revue de linguistique et de didactique des langues* 39, 151-162.
- Blanche-Benveniste, Claire / Valli, André, 1997. *L'Intercompréhension: le cas des langues romanes*, Parigi, EDICEF.
- Bonvino, Elisabetta, 2015. «Intercomprehension Studies in Europe: History, Current Methodology and Future Developments», in: Dolci, Roberto / Tamburri, A. J. (ed.), *Intercomprehension and Plurilingualism: An Asset for Italian Language in the USA*, New York: John D. Calandra Institute Transaction, 29-59.
- Bonvino, Elisabetta / Jamet, Marie-Christine, 2016. *Intercomprensione: lingue, processi e percorsi*, Venezia, Edizioni Ca' Foscari.
- Caddéo, Sandrine / Jamet, Marie-Christine, 2013. *L'intercompréhension: une autre approche pour l'enseignement des langues*, Parigi, Hachette.
- Dijkstra, Ton / van Heuven, Walter J.B., 2018. «Visual Word Recognition in Multilinguals», in: Rueschemeyer, Shirley-Ann / Gaskell, Gareth M., *The Oxford Handbook of Psycholinguistics*, 118-143.
- Gooskens, Charlotte, 2006. «Linguistic and extra-linguistic predictors of Inter-Scandinavian intelligibility», *Linguistics in the Netherlands*, 23(1), 101-113.
- Gooskens, Charlotte, 2013. «Methods of measuring intelligibility of closely related language varieties», in: Bayley, Robert et al. (ed.), *The Oxford handbook of sociolinguistics*, Oxford, Oxford University Press, 195-213.
- Heeringa, Wilbert Jan, 2004. «Measuring dialect pronunciation differences using Levenshtein distance», Groningen, Università di Groningen.

- Heeringa, Wilbert Jan, *et al.* (ed.) 2014. «Modeling intelligibility of written Germanic languages: do we need to distinguish between orthographic stem and affix variation?», *Journal of Germanic Linguistics* 26, 361-394.
- Hilton, Nanna Haug *et al.* 2013. «The influence of non-native morphosyntax on the intelligibility of a closely related language», *Lingua* 137, 1-18
- Kürschner, Sebastian, *et al.* 2008. «Linguistic determinants of the intelligibility of Swedish words among Danes», *International Journal of Humanities and Arts Computing* 2, 83-100.
- Levenshtein, Vladimir I., 1966. «Binary codes capable of correcting deletions, insertions, and reversals», *Cybernetics and Control Theory* 10(8), 707-710.
- Newmark, Peter, 1981. *Approaches to translation*, Oxford, Pergamon Press.
- Otwinowska-Kasztelanic, Agnieszka, 2011. «Awareness and affordances: Multilinguals versus bilinguals and their perceptions of cognates», in: De Angelis, Gessica / Dewaele, Jean-Marc (ed.), *New trends in crosslinguistic influence and multilingualism research*, Bristol, Multilingual Matters, 1-18.
- Ringbom, Håkan, 2007. *Cross-linguistic similarity in foreign language learning*, Bristol, Multilingual Matters 21.
- Schepens, Job / Dijkstra, Ton / Grootjen, Franc, 2012. «Distributions of cognates in Europe as based on Levenshtein distance», *Bilingualism: Language and Cognition* 15(1), 157-166.
- Schüppert, Anja, 2011. *Origin of asymmetry. Mutual intelligibility of spoken Danish and Swedish*, Groningen, Università di Groningen.
- Van Bezooijen, Renee / Gooskens, Charlotte 2005. «How easy is it for speakers of Dutch to understand Frisian and Afrikaans, and why?», *Linguistics in the Netherlands* 22, 13-24.
- Vanhove, Jan / Berthele, Raphaele, 2015. «Item-related determinants of cognate guessing in multilinguals», in: De Angelis, Gessica / Jessner, Ulrike / Kresić, Marijan (ed.), *Crosslinguistic influence and crosslinguistic interaction in multilingual language learning*, Londra, Bloomsbury, 95-118.
- Vanhove, Jan / Berthele, Raphaele, 2017. «Interactions between formal distance and participant-related variables in receptive multilingualism», *International Review of Applied Linguistics in Language Teaching* 55(1), 23-40.

Risorse internet

CoLFIS : *Corpus e lessico di frequenza dell'italiano scritto*,
<<http://esploracolfis.sns.it/EsploraCoLFIS/>>

NVdB : *Il Nuovo vocabolario di base della lingua italiana*,
<<https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana>>