

Nell'universo della didattica delle lingue la valutazione rappresenta uno dei nodi cruciali sia dal punto di vista teorico sia dal punto di vista applicativo, collocandosi al centro della riflessione sull'apprendimento e l'insegnamento delle lingue, tra società, formazione scolastica e formazione universitaria. Il presente volume affronta tale complessità attraverso quattro focalizzazioni che corrispondono ad altrettante sezioni: la prima è dedicata ai principali modelli e idee per la valutazione; la seconda verte sulla valutazione in contesti plurilinguistici e interculturali; la terza è incentrata sul tema del rapporto tra tecnologie e valutazione; la quarta, infine, è intitolata alla valutazione sommativa e ai contesti certificatori.

Francesca Gallina è professoressa associata di Didattica delle lingue moderne all'Università di Pisa. È autrice di numerosi articoli, saggi e monografie, tra cui Italiano lingua di contatto e didattica plurilingue (Franco Cesati, 2021) e Osservare e valutare la competenza lessicale in italiano L2 (Pacini, 2022).

Yahis Martari, professore associato di Didattica delle lingue moderne presso l'Università di Bologna e direttore del CRDI (Centro di Ricerca per la Didattica dell'Italiano), è autore di numerosi saggi in riviste nazionali e internazionali e di alcune monografie. Tra le ultime L'acquisizione dell'italiano da parte di sinofoni (Pàtron, 2017) e Insegnare italiano L2 con i mass media (Carocci, 2019).



€ 32,00

PISA
UNIVERSITY
PRESS

PISA
UNIVERSITY
PRESS

SAGGI E STUDI



DIDATTICA DELLE LINGUE E VALUTAZIONE

DIDATTICA DELLE LINGUE E VALUTAZIONE

Tra società, scuola e università

A cura di Francesca Gallina, Yahis Martari

SAGGI E STUDI

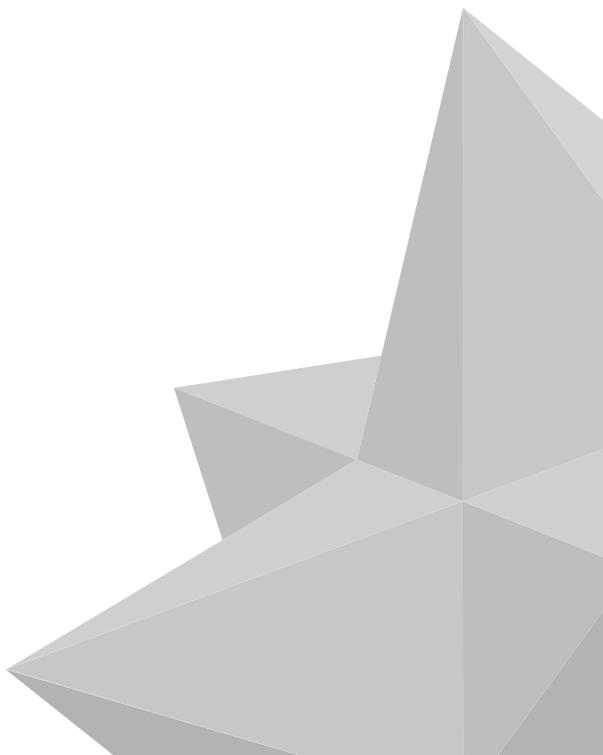


DIDATTICA DELLE LINGUE E VALUTAZIONE

Tra società, scuola e università

A cura di Francesca Gallina, Yahis Martari

PISA
UNIVERSITY
PRESS





SAGGI E STUDI

Didattica delle lingue e valutazione : tra società, scuola e università : a cura di Francesca Gallina, Yahis Martari. - Pisa : Pisa university press, 2023. - (Saggi e studi).

418.0071 (23.)

I. Gallina, Francesca <1978- > II. Martari, Yahis 1. Lingue - Didattica - Valutazione 2. Educazione linguistica - Valutazione

CIP a cura del Sistema bibliotecario dell'Università di Pisa

UPI Opera sottoposta a
peer review secondo
il protocollo UPI
UNIVERSITY
PRESS ITALIANE

Membro Coordinamento
University Press Italiane

Volume pubblicato con il contributo dell'Università di Pisa e della Società di Didattica delle lingue e Linguistica educativa.

Ogni contributo qui pubblicato ha superato una procedura di revisione anonima in doppio cieco (double-blind peer review) che ha coinvolto almeno due revisori scelti tra studiosi ed esperti nel campo della linguistica educativa.

© Copyright 2023

Pisa University Press

Polo editoriale - Centro per l'innovazione e la diffusione della cultura

Università di Pisa

Piazza Torricelli 4 - 56126 Pisa

P. IVA 00286820501 · Codice Fiscale 80003670504

Tel.+39 050 2212056 · Fax +39 050 2212945

E-mail press@unipi.it · PEC cidic@pec.unipi.it

www.pisauniversitypress.it

ISBN 978-88-3339-849-5

In copertina: Foto di Josep Martins su Unsplash.

L'Editore resta a disposizione degli aventi diritto con i quali non è stato possibile comunicare, per le eventuali omissioni o richieste di soggetti o enti che possano vantare dimostrati diritti sulle immagini riprodotte. Le fotocopie per uso personale del lettore possono essere effettuate nei limiti del 15% di ciascun volume/fascicolo di periodico dietro pagamento alla SIAE del compenso previsto dall'art. 68, commi 4 e 5, della legge 22 aprile 1941 n. 633. Le riproduzioni effettuate per finalità di carattere professionale, economico o commerciale o comunque per uso diverso da quello personale possono essere effettuate solo a seguito di specifica autorizzazione rilasciata dagli aventi diritto/dall'editore.

INDICE

Prefazione	
Educazione linguistica e valutazione: dalla classe alla società <i>Elisabetta Bonvino, Matteo Santipolo</i>	7
Introduzione <i>Francesca Gallina, Yahis Martari</i>	11
IDEE E MODELLI PER LA VALUTAZIONE	
Language assessment literacy: overview and way forward <i>Dina Tsagari</i>	19
Assessing CLIL competencies: from theory to practice, a proposal <i>Monica Bezzegato</i>	33
Chi ha paura della soggettività? Studio di caso sulla vaghezza nella valutazione della performance orale in italiano L2 <i>Cristiana Cervini, Yahis Martari</i>	49
Le parole per «imparare ad imparare»: il possesso pluridimensionale del lessico come indicatore di permeabilità alla sua ulteriore espansione <i>Francesca Ferrucci</i>	59
Valutare interazioni orali in italiano L2 con le scale dell'Adeguatezza Funzionale <i>Francesca La Russa, Francesca Pagliara</i>	79
Verso una valutazione della competenza fraseologica per l'Italiano L2 <i>Maria Roccaforte, Luciana Forti</i>	95



Rater Discrepancy Negotiations: How to Develop and Validate a New Scale for Analysis <i>Ece Sevgi-Sole, Aylin Unaldi</i>	105
Valutare l'integrazione di lingua, contenuto e <i>literacies</i> in percorsi CLIL: una sperimentazione nella classe di italiano L2 <i>Fabiana Fazzi, Camilla Spaliviero</i>	117
Specificità dell'arabo e livelli di competenze: problemi e proposte per il livello pre-A1 <i>Cristina Solimando, Alma Salem</i>	133
L'elicitazione orale di frasi relative oblique: strategie di risposta e implicazioni per la didattica <i>Elisa Piccoli, Anna Cardinaletti, Francesca Volpato</i>	149
Cosa correggo? Competenze di valutazione CLIL e insegnanti: tra lingua e disciplina <i>Elisa Corino, Simone Torsani</i>	163
LA VALUTAZIONE IN CONTESTI PLURILINGUISTICI E INTERCULTURALI	
La Valutazione delle Competenze Linguistiche Trasversali: dall'Università alla Scuola <i>Mara Frascarelli, Giorgio Carella</i>	185
L'importanza della formazione linguistica e la valutazione dell'efficacia delle attività pre-partenza (PDO) rivolte ai rifugiati nei programmi di reinsediamento <i>Giulia I. Grosso, Viola Monaci</i>	205
Migranti adulti analfabeti: le competenze predittive per una valutazione prognostica in ingresso <i>Katia Raspollini</i>	219
La rilevazione precoce delle competenze comunicative in contesti plurilingui <i>Maria Chiara Aielli, Marco Mezzadri</i>	233
Testi letterari transculturali e dimensione simbolica dell'intercultura: una sfida per la valutazione? <i>Simona Bartoli-Kucher</i>	247

Rappresentazioni degli aspetti interculturali nei manuali di italiano L2 <i>Antonella Benucci</i>	259
Pragmatica in gioco. Uno strumento per la ricerca e la valutazione formativa <i>Stefania Ferrari</i>	271
“Aiming at a moving target”: un’indagine sull’appropriatezza delle richieste telematiche <i>Victoriya Trubnikova</i>	293

TECNOLOGIE E VALUTAZIONE

Dispositivi e-learning per tracciare e valutare l’apprendimento dell’Italiano L2 in contesti non formali <i>Gerardo Fallani</i>	309
Sviluppare competenze linguistiche attraverso un modello di valutazione formativa automatica <i>Cecilia Fissore, Carla Marello, Marina Marchisio</i>	321
Usabilità e apprendimento dell’italiano L2 online in piattaforme LMS e ambienti aperti: sistemi a confronto <i>Alice Gasparini</i>	333
Proposte per una valutazione formativa per la didattica dell’italiano online <i>Matteo La Grassa</i>	349
Quando si vogliono far valutare al computer anche risposte aperte <i>Carla Marello, Elena Camillo, Marina Marchisio, Cecilia Fissore</i>	361
Valutazione di traduzioni automatiche a reti neurali nella pratica didattica: un’esperienza <i>Mirko Tavosanis</i>	377

VALUTAZIONE SOMMATIVA E CERTIFICAZIONE

Dalla didattica alla certificazione linguistica e ritorno. Come la valutazione certificatoria influisce sull’insegnamento del tedesco oggi <i>Sabrina Bertollo</i>	391
---	-----



La valutazione sommativa dell'italiano a Malta: relazioni tra abilità linguistiche <i>Sandro Caruana</i>	401
Multilingual assessment competences and practices in Europe: ENRICH Erasmus+ <i>Lucilla Lopriore, Kirstin Reed, Dina Tsagari</i>	415
I test di competenza linguistica per scopi accademici. Quali novità per l'italiano L2? <i>Giulia Peri, Adriano Gelo</i>	429
Valutare l'Italiano accademico: problemi e sfide <i>Cristina Pierantozzi</i>	445
L'esame CILS B1 Cittadinanza: validità e affidabilità <i>Aisha Nasimi, Anna Bandini, Laura Sprugnoli</i>	459
Saper progettare e costruire materiali didattici: analisi dei risultati dei candidati all'esame di certificazione glottodidattica DILS-PG di II livello <i>Nicoletta Santeusanio</i>	473
Dalla proposta di un syllabo a un progetto di <i>Scenario Based Assessment (SBA)</i> per valutare l'insegnamento dell'italiano L2 <i>Roberto Tomassetti</i>	493
Temi e sfide della valutazione per la linguistica educativa italiana (e non solo) <i>Francesca Gallina, Yahis Martari</i>	511

Chi ha paura della soggettività? Studio di caso sulla vaghezza nella valutazione della performance orale in italiano L2*

Cristiana Cervini**, Yahis Martari**

1. Domanda di ricerca e introduzione

La domanda di ricerca che guida questo contributo è la seguente: in che modo la logica della vaghezza (o *logica fuzzy*) può aiutarci a descrivere (e accettare) la soggettività e i casi di convenzionalità della valutazione linguistica?

Si noti, innanzitutto, che tale domanda è tutt'altro che rivoluzionaria nel contesto della valutazione: infatti la *logica fuzzy* è in realtà piuttosto rappresentata per i costrutti valutativi, soprattutto nelle scienze matematiche (Hesamian *et al.* 2021; Zhang, Feng 2018) ma anche in quelle umane (ad esempio Boyle *et al.* 2020) e financo nella linguistica (ad esempio Bahi, Necibi 2020; Ivanova, Zlatanov 2019).

La competenza nel linguaggio, così come in altre aree della conoscenza umana, è un *continuum*, perciò si pone sempre il problema di come stabilire i confini dell'«attribuzione di valore» (Dominici 2002). La situazione in cui ci troviamo nel momento in cui dobbiamo misurare con un voto la competenza linguistica di un soggetto – soprattutto nei costrutti basati su contesti comunicativi complessi (Bachman, Palmer 2010) – è talora quella di chi deve stabilire il valore preciso di qualcosa di

* Questo contributo riposa su di una interpretazione dei modelli teorici della *fuzzy logic* presentati e discussi da Moruzzi (2012) con particolare attenzione da parte nostra ai lavori di Timothy Williamson (soprattutto Williamson 1994). La lettura del modello teorico qui presentata costituisce un'evoluzione (seppure nella continuità) rispetto alle idee proposte in Cervini, Martari (2020). Anche se questo lavoro è stato concepito in piena collaborazione tra i due autori, il primo e il terzo paragrafo sono stati redatti da Yahis Martari, il secondo paragrafo da Cristiana Cervini.

** Università di Bologna.



estremamente complesso e sfocato senza avere il tempo o gli strumenti adeguati a una misurazione accurata. Questa stessa situazione si presenta per una certificazione, un *placement test* ma talvolta anche per un voto scolastico o l'esito di uno scrutinio (Shohamy *et al.* 2016). Tale indeterminatezza è del resto rappresentata anche dal QCER – il quale accoglie, per farvi fronte, una dimensione ecologica, interazionista e complessa (North 2022)¹. Si legge ad esempio nel *Volume complementare* (2020: 36):

come i colori dell'arcobaleno, la competenza linguistica è in realtà un continuum [...] nonostante i confini tra i colori siano poco marcati, tendiamo a vedere alcuni colori più di altri». Ma come può funzionare? Come facciamo a capire quando il giallo dell'arcobaleno diventa arancione, o, fuori dalla metafora, quando una competenza di livello B1 diventa B2? Come facciamo a non sbagliare a dire «lì è arancione» oppure «questo apprendente ha un livello B1?

2. Studio di caso

Il nostro studio di caso ha dunque inteso fornire una prima rappresentazione dei fenomeni di vaghezza nel processo di valutazione delle competenze linguistiche. Più precisamente la nostra attenzione si è rivolta alla verbalizzazione esplicita delle ragioni che inducono un valutatore ad attribuire un livello (e non un altro) a una *performance* linguistica.

2.1. Descrizione dell'indagine

In una prima fase dell'indagine, abbiamo coinvolto sei docenti di italiano L2/LS con livelli di esperienza di insegnamento e valutazione eterogenei. Le docenti hanno valutato alcuni *task* preregistrati di produzione orale monologica e di interazione svolti da due studentesse fran-

1. Il *Companion Volume*, fondato su un approccio *action-oriented*, muove, tra gli altri assunti, da un cambiamento «from a paradigm of simplicity (chop things up; don't make things challenging) to a paradigm of complexity: accept complexity and provide reasonable challenges, with scaffolding as necessary» (North 2022: 15).

cofone di italiano lingua straniera (Irène, d'ora in poi I. e Coralie, d'ora in poi C.), con l'obiettivo di attribuire un livello di competenza e di esplicitare le ragioni e i fattori alla base della decisione presa². A disposizione delle valutatrici vi erano i file audio-video, da riguardare a piacere e senza vincoli di tempo, la griglia dei descrittori del Companion Volume (2018: 171-172) e un questionario con sette gruppi di domande a risposta aperta e chiusa. Attraverso il questionario abbiamo inteso comparare i) i criteri di valutazione adottati e percepiti come discriminanti da parte delle docenti valutatrici su medesimi compiti (es.: Sulla base di quanto riportato globalmente nei descrittori del QCER alle voci correttezza, fluenza, estensione, interazione, coerenza, assegna un livello alle produzioni di I. e C.; Per ogni descrittore, puoi spiegare le ragioni – almeno tre – che stanno alla base della decisione che hai preso? Puoi indicare esempi positivi – usi corretti – e negativi – usi impropri – presi dalle produzioni orali?); ii) gli elementi linguistico-comunicativi delle performance orali che avrebbero determinato una decisione valutativa differente rispetto a quella presa (es.: Per classificare nel livello immediatamente superiore/inferiore le performance di C. e I., cosa avresti dovuto osservare?); iii) il grado di certezza o incertezza associato alle valutazioni espresse (es.: Quanto ti senti certo/a del livello assegnato alle prove?); iv) l'utilità percepita nei confronti dei descrittori del QCER (es.: I descrittori del QCER ti sono stati utili per valutare le produzioni orali?).

2.2. Argomentazione dei criteri

Riprendendo la metafora dei colori, in cui ciascun colore corrisponde a un tratto della competenza linguistico-comunicativa, ci siamo chiesti se alcuni colori/tratti spicchino più di altri nelle argomentazioni delle valutatrici e se queste differenze incidano negativamente sul loro grado

2. I video sono tratti dal progetto *Productions orales illustrant les 6 niveaux du Cadre européen commun de référence pour les langues* elaborato nel 2008 dal CIEP (Centre International d'études pédagogiques, ora France Education International), in collaborazione con il Consiglio d'Europa (Divisione Politiche linguistiche), Cambridge ESOL, Istituto Cervantes, Goethe Institut, Università di Perugia. Per maggiori informazioni, cfr. <https://www.france-education-international.fr/productions-orales-illustrant-les-6-niveaux-du-cecr/?langue=fr> (ultima consultazione, ottobre 2022).

di accordo. A questo proposito, è interessante notare che, all'unanimità, tutte le valutatrici hanno associato la performance orale di C. al livello B1 del QCER, adducendo tuttavia ragioni che provengono da prospettive diverse: il criterio *correttezza* è argomentato talvolta in riferimento a deviazioni dalla norma, talvolta a usi virtuosi della lingua. Gli errori sono descritti come «sistematici», «di base», «sporadici», «interferenze dalla L1», «fossilizzazioni» e associati a esempi eterogenei di tipo lessico-grammaticale (presenza dell'articolo davanti all'aggettivo possessivo, assenza di accordo di genere e numero, uso di «avione» in veste di calco errato dal francese al posto di «aereo»). Per quanto riguarda gli usi virtuosi della lingua italiana, si menzionano la padronanza nel distinguere il passato prossimo dall'imperfetto, l'uso di collocazioni ed espressioni idiomatiche, la presenza di costruzioni sintattiche articolate, l'uso di aggettivi precisi e lessicalmente raffinati come nel caso di «talentuoso». Un andamento simile nell'elaborazione delle risposte è osservabile per il criterio *fluenza*: per alcune valutatrici le pause durante l'eloquio sono «troppo frequenti e troppo lunghe», per altre invece «la fluenza è ottima, con poche pause e pochi tentennamenti». In alcuni casi ancora la fluenza viene argomentata in riferimento alla «comprensibilità del discorso e alle capacità di riformulazione», senza rivolgere attenzione specifica alla qualità delle pause. La cura delle scelte lessicali viene messa in evidenza dalle valutatrici anche nel criterio estensione, in cui però si mescolano riflessioni su aspetti molto eterogenei della competenza linguistica: esitazioni, lunghezza dei turni, velocità dell'eloquio, lunghezza delle frasi. È evidente quindi che alcuni «colori», simbolo dei tratti della competenza linguistico-comunicativa, colpiscono l'attenzione del valutatore con minore o maggiore intensità, portando ad argomentazioni talvolta anche sostanzialmente diverse. Eppure, ciononostante, le valutatrici convergono verso le medesime decisioni, come da sintesi riportata in tabella 1.

Tabella 1. Sintesi livelli prima indagine.	
Studentessa	Livello assegnato dalle valutatrici
Coralie	B1: 100% (6 su 6)
Irène	B1: 50% (3 su 6)
	A2: 50% (3 su 6)

2.3. Utilità del QCER

È abbastanza intuitivo immaginare che la condivisione di medesimi descrittori, quali sono quelli offerti nel QCER, abbia agevolato l'armonizzazione delle valutazioni finali, in modo simile a quanto avviene nelle sessioni di standard setting. Rispetto all'utilità, i descrittori per la valutazione del parlato del QCER sono percepiti come «molto utili» o «utilissimi» da cinque valutatrici su sei, e «abbastanza utili» da una valutatrice. Leggiamo per esempio: «Grazie agli indicatori è possibile valutare più attentamente i vari aspetti dell'oralità, sebbene spesso i confini tra i livelli non risultino così marcati e semplici da interpretare. In particolare, nel task 2 ho trovato alcune difficoltà nel giungere a un giudizio complessivo in quanto ho notato nella studentessa da osservare uno squilibrio tra le varie componenti, in particolare fluenza e correttezza/estensione».

Questa argomentazione ci sembra particolarmente significativa perché coerente con le nostre riflessioni sui confini di attribuzione di valore. In un altro caso, i descrittori, seppur reputati molto utili, risultano per la valutatrice troppo vaghi, se non associati a strumenti più analitici come i sillabi di italiano L2.

2.4. Ampliamento dell'indagine

È noto come dietro a ogni processo di valutazione e ad ogni decisione valutativa ci sia una concezione della lingua e del comunicare. McNamara (2000: 3) affermava «testing is a universal feature of social life». Infatti, in prospettiva diacronica, osserviamo che i modelli di valutazione si trasformano parallelamente all'evoluzione degli approcci didattici: dai test discreti/fattoriali, ai test integrati, fino a quelli comunicativi, e alla valorizzazione degli scenari azionali basati su principi di autenticità situazionale e interazionale. Di fronte a questa consapevolezza, ci siamo chiesti se un accordo così elevato ottenuto nell'assegnazione dei livelli da parte delle sei valutatrici fosse influenzato dalla loro cultura didattico-valutativa di provenienza. Abbiamo dunque esteso l'indagine a un gruppo di altre quattro valutatrici, in questo caso serbe e docenti di italiano LS a Belgrado. L'indagine è stata condotta esattamente con gli stessi strumenti e le stesse modalità: video, questionario, descrittori

del QCER. I risultati ottenuti sembrano confermare la fondatezza della nostra intuizione. Complessivamente, infatti, tutte le valutatrici serbe hanno assegnato il livello A2 alle performance di Coralie, dunque inferiore rispetto alle valutatrici italiane, tutte concordi sull'assegnazione di un livello B1. Osserviamo inoltre che, come nella prima indagine, la valutazione della performance di Irène ha evidenziato una partizione in due del campione di risposte. Sia nel contesto italofono che in quello serbo, la valutazione della performance di Irène risulta più sfocata.

Tabella 2. Sintesi livelli seconda indagine.	
Livello performance Coralie	A2: 100% (4 su 4)
Livello performance Irène	B1: 50% (2 su 4) A2: 50% (2 su 4)

Nelle singole argomentazioni dei criteri, permane una certa eterogeneità nella descrizione delle caratteristiche del parlato in interlingua, seppur tutte le valutatrici attribuiscono notevole valore agli errori grammaticali, soprattutto di accordo di genere e numero. Salta agli occhi anche la differenza nella percezione dell'utilità del QCER, verso la cui utilità le valutatrici serbe dimostrano meno entusiasmo (i descrittori sono giudicati «utili» e non «molto utili» o «utilissimi»).

Questi primi risultati, osservati per ora su un numero molto esiguo di partecipanti, suggeriscono di approfondire le dimensioni di convenzionalità e di vaghezza nella valutazione, anche in riferimento al contesto didattico-valutativo e culturale in cui si opera e all'abitudine, più o meno diffusa, di avvalersi dei descrittori di riferimento del QCER.

3. Interpretazione. *Fuzzy logic* per la valutazione linguistica

3.1. Vaghezza e condizioni di sicurezza

La teoria epistemicista (Williamson 1994) definisce la vaghezza come «conoscenza inesatta», il che non significa insufficiente alle attività della

vita, ma inevitabilmente caratterizzata da confini sfocati³. La vaghezza non rende arbitrario l'atto della valutazione, il limite tra un voto e l'altro, ma solo inevitabilmente basato su di una credenza plausibile e sufficiente, in virtù di una «condizione di sicurezza».

L'esempio noto è il seguente. Se ci venisse chiesto di dire quanti grani ci sono in un mucchio di sale, e non potessimo contarli, faremmo una stima e stabiliremmo un intervallo (**I**) che è certamente *plausibile*. Se dovessimo dire però qual è il numero esatto (**x**) dovremmo cercare di indovinare scegliendo in (**I**). **X** sarebbe solo una credenza e potrebbe essere vero oppure no. Ma fuori da **I** sarebbe implausibile. La condizione di sicurezza di **x** (cioè la sua accettabilità) è *che tale credenza* «non potrebbe facilmente essere falsa». La condizione di sicurezza si restringe man mano che la nostra conoscenza si fa più esatta: nel nostro esempio, avendo a disposizione il tempo per contare più o meno precisamente i grani.

Questo stesso concetto di «condizione di sicurezza» è quello che noi applichiamo alla valutazione di un soggetto e della sua competenza linguistica, soprattutto nel caso di una valutazione in cui la nostra conoscenza sia inesatta e la valutazione soggettiva. Cioè tutte le volte in cui non possiamo misurare *tutta* la conoscenza linguistica di un soggetto. Indichiamo una **I** (un giudizio) e in alcuni casi, se dobbiamo indicare una **x** (un voto), sarà un valore che «non può facilmente essere falso».

Ad esempio, nel nostro primo campione, tre soggetti valutano B1 (**I**) un apprendente rilevando questioni diverse:

Nonostante qualche errore sistematico di morfologia (*il mio padre, la mia madre; i suoi ruoli*) e qualche sporadica interferenza lessicale dallo spagnolo (*semana, naturaleza*), complessivamente il repertorio associato a situazioni di routine appare corretto. In particolare, risulta efficace l'opposizione imperfetto – presente per raccontare la trama del film (anche se un passaggio non risulta comprensibile), impeccabile l'utilizzo dei pronomi diretti (*può non vederla, il ragazzo lo fa*) e della preposizione *da* con valore temporale.

3. Per una discussione sulla letteratura scientifica sull'argomento rimandiamo a Cervini, Martari (2020).



Fraasi come «mi piacerebbe andare in Italia», «così possiamo prendere il sole», usate durante la conversazione, oppure l'uso corretto di *ciòè* durante la descrizione del film nel monologo. Ci sono però fossilizzazioni in alcune forme tipiche di livelli inferiori evidenti in frasi come «la mia madre è francese», «mio padre è di Perù» oppure durante il dialogo «le vacanze sono per rilassare».

Possiede un linguaggio sufficiente per esprimersi (*prendere il sole, andare al mare, ecc*) ma fa sistematicamente errori di base (*nell'estate/ la mia madre/ i suoi ruoli al posto di i loro ruoli*), utilizzo appropriato dell'aggettivo talentuoso.

3.2. Margine dell'errore, la convenzionalità, zone borderline

I confini di I sono sfumati e quindi ambigui. Così, sempre secondo Williamson (1994), esiste un *principio di margine dell'errore*: le nostre credenze non sono abbastanza affidabili da permetterci di definire il confine in modo univoco e dobbiamo accettare casi *borderline*, rispetto ai quali, oltre a non avere una conoscenza esatta, occorre che la nostra credenza sia considerata plastica. Esiste dunque una situazione di Principio di margine dell'errore anche in senso metalinguistico (Moruzzi 2012: 238) che fa sì che la possibilità di scarto tra vero e falso, tra un livello e l'altro (I e non-I) unicamente in base a due valori di x sia una credenza non accettabile. E così non sarebbe ragionevole ad esempio in base alla mancata conoscenza di un solo item lessicale decidere che un soggetto sia B1 invece che B2.

Più è preciso il TEST, più i valori di I sono limitati (A1++). Ma questo richiede spesso un accordo «convenzionale», cioè concordato da una comunità. Come avviene per definire quando una collina smette di essere definibile collina e inizia a essere una montagna (600 s.l.m.), e come avviene, ad esempio, alla fine di uno *standard setting*.

Nel contesto molto specifico del *placement test*, per fare un altro esempio, il lato convenzionale è maggiormente presente e accettabile, dal momento che tutta l'attività di misurazione è realizzata rispetto a un contesto molto specifico (quello delle risorse e della struttura dell'istituzione che eroga corsi di lingua). In tutti questi casi una comunità decide che la distanza tra x e $x+1$ è discriminante, appunto in modo convenzionale.

Fuori dal contesto della convenzionalità, invece, è possibile che alcuni valori di x che si trovano nella zona *borderline* siano complessivamente interpretati come appartenenti a I oppure come non appartenenti a I. Ci sono zone *borderline* che fanno sì vi sia una variazione di I che può essere però unicamente I+1 o (cioè non può essere I+2). Ad esempio, nel nostro primo campione, Irène è valutata B1 o A2 con questi pressoché identici giudizi di correttezza:

Possiede un linguaggio non abbastanza sufficiente per esprimersi (tutto il dialogo ruota attorno al termine francese *randomnées* che ovviamente le due parlanti capiscono essendo francofone). Fa sistematicamente errori di base (*il mio padre/ a me non mi piace molto, nel perù*)

Numerose incertezze nella morfologia verbale (*ha frentato/ frontato > ha affrontato*; incertezze nella coniugazione del verbo preferire); *il mio padre*; ha usato per 3 volte “*randomnées*” e un calco dal francese “*alberghi di gioventù*”; nell’*avione* (aereo).

Tale oscillazione può riguardare i singoli valutatori ma anche le culture didattico-valutative. Così osserviamo, nel secondo campione, un allineamento delle valutatrici su di un altro giudizio, rispetto alle valutatrici del primo campione. In sintesi, la valutazione linguistica indica per lo più delle stime ragionevoli (dei valori I): cioè, definisce intervalli di competenza linguistica (A, B, C; oppure A1, A2, B1 etc.) in base a valori adiacenti: sbagliare a valutare uno di questi valori non è quasi mai sufficiente a non concordare su I. Per tale condizione di sicurezza, così come altre cose, la valutazione linguistica, in rapporto alle convenzioni di un sistema valutativo, funziona.

Bibliografia

- Bachman L.F, Palmer A.S., 2010, *Language assessment in practice*, Oxford, Oxford University Press.
- Bahi H., Necibi K., 2020, *Fuzzy Logic Applied for Pronunciation Assessment*, in «IJCALLT, IGI Global», 10, 1, pp. 60-72.



- Boyle B., Mitchell R., McDonnell A., Sharma N., Biswas K., Nicholas S., 2020, *Overcoming the challenge of “fuzzy” assessment and feedback*, in «Education + Training», 62, 5, pp. 505-519.
- Cervini C., Martari Y., 2020, *Soggettività, vaghezza e ambiguità nel language assessment. Note sui confini dell’attribuzione di valore*, in «Quaderni di semantica», Special issue, pp. 101-130.
- Consiglio d’Europa, 2018, *Common European framework of reference for languages (CEFR). Learning, teaching, assessment. Companion Volume with New Descriptors* (disponibile online in <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>), Ed. it. *Quadro comune europeo di riferimento per le lingue. Apprendimento insegnamento valutazione. Volume complementare con nuovi descrittori*, Trad. it. di M. Barsi e E. Lugarini, Università degli Studi di Milano, Italiano LinguaDue, 2020, 2.
- Domenici G., 2002, *Manuale della valutazione scolastica*, Roma-Bari, Laterza.
- Hesamian Gh., Akbari M.G., Shams M., 2021, *Fuzzy Assessment of Heavy Metal Pollution*, in «International Journal of Industrial Mathematics», 13, 4, pp. 385-394.
- Ivanova V., Zlatanov B., 2019, *Application of fuzzy logic in online test evaluation in English as a foreign language at university level*, AIP Conference Proceedings 2172, 040009.
- McNamara T., 2000, *Language Testing*, Oxford, Oxford University Press.
- Moruzzi S., 2012, *Vaghezza. Confini, cumuli e paradossi*, Roma-Bari, Laterza.
- North B., 2022, *Companion Volume and the action-oriented approach*, in «Italiano LinguaDue», 2, pp. 1-23.
- Shohamy E., Or I., May S. (eds.), 2017, *Language Testing and Assessment*, New York, Springer.
- Williamson T., 1994, *Vagueness*, London-New York, Routledge.
- Zhang P., Feng G., 2018, *Application of fuzzy comprehensive evaluation to evaluate the effect of water flooding development*, in «J. Petrol. Explor. Prod. Technol.», 8, pp. 1455-1463.