



**HAL**  
open science

## Evaluation and plurilingual competence

Emmanuelle Huver

► **To cite this version:**

Emmanuelle Huver. Evaluation and plurilingual competence. Silvia Melo-Pfeifer; Christian Ollivier. Assessment of Plurilingual Competence and Plurilingual Learners in Educational Settings, Routledge, 2023, 9781003177197. 10.4324/9781003177197 . hal-04885709

**HAL Id: hal-04885709**

**<https://hal.science/hal-04885709v1>**

Submitted on 14 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Evaluation and plurilingual competence: between epistemological consistency and ethical vigilance

Emmanuelle Huver

University of Tours, EA 4428 Dynadiv

Questions about the evaluation of plurilingual competence have always accompanied reflections on plurilingual competence itself, and this has been the case from the start of work on this topic. Therefore, in the research literature on these issues, it appears from the very first reflections on the subject, included in the French-language literature, on which I will preferentially rely here<sup>1</sup>.

These different works regularly highlight the difficulty, nay the impossibility, of assessing plurilingual competence: "assessment" and "plurilingual competence" are then presented as belonging to two paradigms that are so different that their combination would practically be an oxymoron. This tension between the need to assess plurilingual competence and the (technical and/or ethical) impossibility of doing so is still present today, including in the argument of this book. The ways of trying to resolve this contradiction have been numerous. They have even multiplied as the notion of plurilingual competence was growing in importance, particularly at the institutional level: the publication of the *Companion Volume of the Common European Framework of Reference* (Council of Europe 2018) can be considered as one of the most recent avatars of that process.

In this chapter<sup>2</sup>, I will argue that the research on plurilingual competence assessment has essentially consisted of a reflection on assessment tools and approaches, and, logically, on designing such tools and approaches. By doing this, the research has neglected a more theoretical / epistemological reflection on the tensions mentioned above between assessment and plurilingualism, that brings up, more broadly, the problem of the status and conceptions of the notion of diversity at the basis of these tools.

I don't have enough space here to present extensively these different conceptions, but I can sum them up roughly by opposing two conceptions of that notion. On the one hand, diversity can be understood as a *quantitative* notion, in which case language diversity, for example, refers to the fact that languages can be objectively distinguished, enumerated, cumulated. On the other hand, diversity can be understood as a *qualitative* notion that questions of the very notion of language as a homogeneous system with absolute and predefined contours. In that conception of diversity, what matters is what's happening in the *relation* between languages (i.e. interlinguistic and translinguistic processes, but also biographical aspects when a person puts into relation different "languages" that crossed ones life) and the experience of otherness that results from that linking<sup>3</sup>.

I will argue that this until now quite neglected reflection is yet essential to understand what can be diversely meant by "assessing plurilingual competence", and thus to define conceptions of plurilingual competence and evaluation that *both* are based on (qualitative) diversity as a core value whether in the engineering, political or ethical field. To this end, I will first provide an overview of recent research on assessment in language didactics, paying close attention to work on plurilingual competence. Then, I will show how monolingual assessments and the research criticizing them are based on a conception of language and assessment that is, in the end, quite convergent. This will allow me, in the last part, to make some proposals—epistemological in nature—as to what could be a languageS assessment that takes diversity seriously.

---

<sup>1</sup> This language choice lies on the fact that, on the one hand it is my main field of reflection and that, on the other hand, it seems important to me that these elements, which are often little known or little cited by the English-language research literature, can be disseminated.

<sup>2</sup> I hereby thank Katia Schuchman, who took in charge the translation of this text. Of course, I am solely responsible for the content and the possibly remaining mistakes.

<sup>3</sup> For more details on this matter, see, among others, Coste (2002), the work of V. Castellotti and D. Moore (for example Castellotti and Moore 2011) or Huver (2014).

# 1. Recent work on evaluation in language teaching: between criticism and engineering

To sum up roughly, most of the work on evaluation in language didactics can be divided into two main trends: one that we could “critical” on the one hand, and one that we could call “engineering” on the other hand<sup>4</sup>.

**The critical trend** examines the political and ideological stakes on evaluation. If we stick to the most recent works, it consists essentially in a criticism of certifications, or at least of their prevalence. Researchers denounce the fact that the predominance of certifications leads to an overvaluation of descriptors and standards, with the corollary of a homogenization of the proposed evaluations, an excessive influence of certification bodies on language teaching/learning, and a reduction of educational stakes to their market logic. Within this "critical" theoretical framework, one frequently finds the idea that evaluation, or more precisely, standardized evaluation, would be at the service not of education but of a capitalist and managerial ideology (cf. for example Puren and Maurer 2019). Moreover, these analyses converge, although they are not actually being treated as linked, with the work of "critical" sociolinguistics, which considers languages (and in particular linguistic diversity) as a capital inscribed in power relationships, or even participating in types of domination by capitalism (cf. the work of M. Heller, A. Duchêne or, from a slightly different angle, Ph. Blanchet).

These criticisms of evaluation, and more particularly of certification, are part of a larger criticism of the ideology of evaluation, from an economic and ideological perspective inspired by Bourdieu's critical sociology. They simultaneously draw on reflections from other fields of the humanities (philosophy, sociology, political science, etc.) and, for language didactics, on work from English research (e.g. Spolsky 1995, Shohamy 2001, Mc Namara and Roever 2006, etc.). In this framework, assessment (in fact, more precisely, standardized assessment) is seen as a "social technology" (Madaus 1990) used to support a more or less explicit ideology of selection, which (among others) deeply question the goals of compulsory schooling.

The **engineering trend** is essentially interested in the building of tools and approaches (and, more secondarily, on practices) to enhance relevance to a specific chosen teaching methodology (communicative, action-oriented, task based, etc.). This work can take the form of methodological guides drafting evaluation tools, guides that are either general (Bachman and Palmer 2010) or targeted at specific didactic areas or audiences (Kunnan, ed. 2014, which is intended to be a reference work in the field). Most of them aim to promote alternative assessment approaches to certification and, more widely, to summative assessment and to the use of numerical marks. The aim then is to foster assessment practices:

- for language learning and/or learning to learn languages (Médioni 2016, Cuq and Davin 2016, Soubre 2021)
- consistent with the latest methodological approaches promoted, foremost among which is the action-oriented perspective widely disseminated by the CEFR.

We can observe a real enthusiasm for this trend in the language testing field, as evidenced by the increase in titles published on the subject over the past five years (for example and without being exhaustive: Cheng and Fox 2017, Davis et al. 2018, Hidri ed. 2018, Shohamy et al. 2017, Taguchi et al. 2018, etc.). Most of the tools developed may have in common to promote “alternative” (i.e. more positive and formative) assessment, but they do not really address explicitly (or rarely) the issue of assessing plurilingual competence, even though there is a growing interest in this topic. Moreover, they mostly stick to the observable dimensions of the competence only, without questioning the didactical and political consequences of this choice that finally avoids the “inaccessible” dimensions of assessment process and the subjectivity that it would introduce in evaluation (Huver, 2014b).

---

<sup>4</sup> These two trends are of course not completely mutually exclusive, even if the critical tendency is undoubtedly more recent. This one is based indeed on multidisciplinary works on the ideology of evaluation published over the last fifteen years as a part of a more general movement denouncing the excesses of neo-liberalism.

## 2. What about research on plurilingual competence assessment?

### 2.1. Similar structuring

The two trends highlighted in the field of languages assessment are also reflected in works on the assessment of plurilingual competence.

**The critical branch** consists essentially of a criticism of monolingual evaluations on the grounds that:

- they do not allow for the entire language repertoire of individuals to be taken into account, even in situations where this is, didactically and/or politically, particularly relevant. For example, the people's orientation at the start of their training would benefit from being relying not only on their "level" in the target language, but on a range of information concerning their language biography (cf. in France, the assessments carried for the orientation of migrant pupils at school, which include an assessment in their so-called original language).
- they are often based on a very normative conception of language: people are then evaluated in relation to a fantasized "native speaker", who would speak a "pure" language or, at least, a very widely homogenized one (i.e., redacted from its variations). According to the proponents of this type of criticism, the consequences of this normativity and homogenizing conception of language are both didactic and political. On the didactic level, monolingual assessments are denounced as essentially subtractive (people are assessed on what is *not* conform to the dominant norm). On the political level, they are denounced for ignoring the identity dimensions of language, resulting in stigmatizing or discriminatory effects.

These works criticizing monolingual assessments are quite diverse. Some of them deny descriptors and scales per se (often by referring at Bourdieu's critical sociology) or criticize the CEFR as based on irenic plurilingualism. Some say it evacuates subjectivity and subject (Prieur and Volle 2016, Anderson 1999, Anquetil et al. 2017); for others it does not mention the issues of commodification of education (Maurer and Puren 2019); for others still it lacks of a fundamental epistemological reflection on the notion of diversity, that would connect the above-mentioned other criticisms (Castellotti 2017, Huver 2014). But some others recognize the CEFR, including descriptors and scales, as an attempt to overcome the homogenization of language, or at least the systematic reference to the so-called "native speaker" in assessments (cf. arguments in favour of the *CEFR Companion Volume* in Piccardo et al. 2019).

However, in the field of plurilingual competence assessment, **the engineering branch** is also the most developed (like in language assessment in general). A great deal of research has been dedicated to the preparation and testing of tools and approaches to assess plurilingual competence, including:

- tools for the evaluation of the various approaches that fall within the so-called "plural approaches" (which include: intercultural approach, integrated didactics, intercomprehension, language awareness). Thus, for example, articles or book chapters are devoted to the evaluation of intercomprehension (e.g.: Ollivier 2019);
- tools directly concerned with plurilingual competence. These tools can either be reference systems incorporating descriptors, such as the *Framework of reference for pluralistic approaches to languages and cultures*, the Harnos project<sup>5</sup> or the *CEFR Companion Volume*, or tools for assessing plurilingual competence itself. In this regard, we can mention in particular, in addition to the *European Language Portfolio*, the proposals of Lenz and Berthele (2010) and Goullier (2017: 266ff.), the experimental plurilingual assessment scenarios of the Val d'Aosta (linked on specific curricula), and the mediated language tasks of the Inca project (Huver and Springer 2011).

While many of these tools explicitly follow the guidelines of the CEFR, others distance themselves from it, but still claim in the same time:

---

<sup>5</sup> Harnos is a Swiss intercantonal agreement on the harmonization of compulsory schooling between the different Swiss cantons, approved by referendum in 2006. Among other things, it includes national educational standards in various school subjects, including foreign languages and schooling. For more details, see Lenz (2006).

"an 'integrated assessment' which takes into account all the issues at stake in the language teaching-learning process; an 'integrated plurilingual methodology' which relies on the already existing in terms of learners' language repertoire and on the already constructed in terms of knowledge about languages and language learning skills" (Maurer and Puren 2019).

Research on the assessment of plurilingual competence is therefore a complex and heterogeneous field, with its areas of dissensus and internal tensions. Nevertheless, however diverse they may be, these studies agree on some common principles. In particular, recommendations that transversely aim denounce the (formative *and* social) selection linked to assessment, the negative effects on pupil's well being, learning capacity and empowerment. Reciprocally, they promote assessments that would be less "threatening" (Butera et al. 2011), less "destructive" (Shohamy 2001), less "glottophobic" (Blanchet 2016), and therefore more qualitative (not based on a score, a fortiori graded), more formative (strengthening learning), and more inclusive (of people's linguistic and cultural diversity especially) — on this point, cf. Goullier (2017: 260), or Cavalli in this book. This has played a major role in raising awareness of the political and even identity-related issues associated with evaluation. It has also contributed to the diversification of evaluation tools and to new ways of dealing with the linguistic and cultural diversity (cf. the emblematic example of the *European Language Portfolio*).

The other common point is that even the critical works consider the interventionist aim of language didactics as essential, so that they propose, as a counterpoint to their criticism, other *tools*. But doing so, they also reinforce in the end the focus of language didactics research on the engineering dimension.

What are the implications of these findings for research on the assessment of plurilingual competence?

## 2.2. Plurilingual assessment, truly?

This focus on the engineering aspects of language didactics ultimately leads to the idea that exercises, activities and evaluation tools are *in themselves* destructive (or constructive), threatening (or benevolent), glottophobic (or glottophilic), depending on whether they are closed (or open), quantitative/graded (or qualitative/ungraded), monolingual (or plurilingual), and exclusive of diversity and variation (or inclusive of these). This contributes to evacuating from researches representations on the one hand and epistemological questions on the other. However, without this epistemological questioning, the conception and the status of diversity are notoriously excluded from reflection:

- first of all, because the notion of plurilingualism (and, more broadly, of linguistic plurality/diversity) is not defined, the work of the years 1990-2000 on the subject being generally considered sufficient<sup>6</sup>;
- secondly, because linguistic diversity is maintained as an *object* to be assessed, based on the observable and objective dimensions of it.

Thus, from this double point of view, there is paradoxically a form of (epistemological) convergence between assessments that are denounced as being monolingual and those that claim to be plurilingual. In monolingual assessments indeed, (linguistic) diversity is certainly absent; but in so-called "plurilingual" assessments, it is generally reduced to quantitative forms of diversity:

- basically, a tool is offered in X languages or includes X languages
- they rarely include trans- and inter-linguistic dimensions. And even if the trend is growing (Lenz and Berthele 2010, Chardenet 2005, Council of Europe 2019, etc.), the cited researches reduce plurilingualism to observable dimensions of the plurilingual competence (vs question the stakes of including the more unobservable dimensions of the plurilingual *experience*);
- and even more rarely, they try to seize the ensemble of issues that question evaluation *from the perspective of* diversity. In other words, they do not try to imagine what could be an evaluation in which diversity would not only be the object but mainly the core value (cf. recently, nevertheless, Gajo et al. 2019). The next part of this article will try to develop more

---

<sup>6</sup> The newly introduced notions of "translanguaging" or "superdiversity" do not qualitatively change this reflection.

precisely what I mean by “considering diversity as core value” in the assessment field (i.e. what I have also called “a methodical use of diversity” or a “diversitarian” evaluation<sup>7</sup>).

The same rationale can be extended to conceptions of evaluation. Indeed, even when they include elements of diversity (and/or when they claim to assess plurilingual competence), the different assessments and/or reflections on assessment eventually remain epistemologically homogeneous by the fact that they:

- generally maintain the idea that some assessment tools or practices would be inherently better than others;
- conceptualize the intrinsically interpretive (hermeneutic) nature of evaluation only to a limited extent, except in an attempt to “control” or “master” the assessment process, that is then considered as (too) subjective (see below);
- Thus, they perpetuate a “technical thought” in languages assessment, beyond the formally qualitative aspects that are yet considered inseparable from the assessment of plurilingual competence (on the question of “technical thought” in language didactics, see Debono 2020).

In the end, since diversity remains the *object* of assessment but not its *principle*, language assessment can be considered to remain ultimately monolingual deep down, including the assessment of plurilingual competence. In other words, the “reversal” promised by the introduction of plurilingual competence in language didactics (Coste 2002) does not seem to have affected assessment, if not very superficially:

Un système de gradation qui a été d’abord pensé et produit en relation à une langue reste inscrit dans une logique monolingue de progrès continu apprécié tout au long selon les mêmes critères. Un tel système ne peut tout simplement pas accommoder la pluralité linguistique et culturelle dans ce qu’elle a – justement – de plurilingue et dans les discontinuités, les recompositions, les modifications d’équilibre entre les différentes composantes du répertoire que présente l’évolution « ordinaire » des capacités plurilingues et pluriculturelles de tout acteur social. (Coste, 2021)<sup>8</sup>

### 3. Taking diversity seriously in languages assessment

Then, what can it mean to let language didactics (and languages assessment) be affected (in the strongest sense of the word) by plurilingual competence, or more precisely by the principle of diversity? In other words, what can it mean to “take diversity seriously” in languages assessment and thus put diversity at the heart of “language” and “assessment”?

#### 3.1. Language(s): thinking on the basis of linguistic diversity

As explained above, assessing in a plurilingual manner is not simply a matter of assessing plurilingual competence, let alone as a juxtaposed assessment of different languages, based on level descriptors designed a priori.

In my work and in the work of my research team (Dynadiv, <https://dynadiv.univ-tours.fr/>), the notion of language is not initially conceptualized as a system or a means (of communication), nor is it considered primarily as a heterogeneous social phenomenon, but as ways of experiencing the world. Let us also specify that, in this orientation, an experience is not only individual and, consequently, does not evacuate the political dimension: the notion of experience as conceptualized here takes on a historical thickness and density and, in doing so, a collective dimension requiring a political treatment (cf. in this regard Romano 2010).

---

<sup>7</sup> See also a number of my works over the past ten years).

<sup>8</sup> “A system of gradation that was first conceived and produced in relation to *one* language remains inscribed in a monolingual logic of continuous progress appreciated throughout according to the same criteria. Such a system simply cannot accommodate linguistic and cultural plurality *in its plurilingual aspect* and in the discontinuities, the recompositions, the modifications of balance between the different components of the repertoire that the “ordinary” evolution of the plurilingual and pluricultural capacities of any social actor presents.”



Furthermore, it is worth asking what "it means to look at a language as a foreign language" (Gajo et al. 2019). According to them—and I largely agree with them:

"The answer immediately involves the notions of contact and context: a language is foreign because it comes into contact with at least one other language, either in a particular social context or in the trajectory of a learner."

At the didactic level, the intersection of these two reflections (language as an experience of the world/foreign language learning as an experience of contact and otherness) implies considering that learning/teaching a language is first and foremost a way of experiencing otherness:

"Learning/teaching a language is to (let) experience and confront a certain way of being in the world; and consequently, learning/teaching a *foreign* language is to (let) experience and confront *another* way of being in the world, *another* way of being human" (Huver and Macaire, 2021).

However, in existing assessments, the underlying conception of language is always either structural or communicative/task-based, and therefore pragmatic. This also applies to assessments of plurilingual competence (cf. the tools mentioned above, which are mostly based on the methodological principle of authentic communicative tasks).

Moreover, the question of language contacts is never really raised, except at worst under the angle of interference (to be banned) or, at best, under the angle of error (to be explained). However, from my approach, language evaluation is necessarily languages evaluation (e.g. is always plurilingual). In other words, it cannot avoid the question of contacts, considering them not only as normal, but also as a certain "way of being" (to use Beacco's terms about plurilingualism) and as part of what Castellotti (2017) calls "appropriation"<sup>9</sup>.

Assessing plurilingual competence from a pragmatist perspective is, in the end, quite easy—even if it involves a sophisticated didactic technology—since, reduced to its semiotic dimensions, language is more easily predictable, and therefore controllable, and therefore assessable. The situation is different when language learning is considered from the perspective of appropriation. Indeed, in this framework, although the evaluation of plurilingual competence may<sup>10</sup> still be possible, it can no longer be done through standards and descriptors designed a priori, and according to the verticality of levels. In other words, working on the notion of language through the prism of the principle of diversity also implies working on the notion of evaluation through this same prism.

### 3.2. Evaluating: understanding and interpreting

As we have already said, the question of assessing plurilingual competence is often asked in the light of the difficulties inherent in this act, or even in terms of its impossibility (Huver 2010). The *relevance* of this assessment is less often mentioned, although one might ask, with de Pietro (2019):

"If the strictly speaking *plurilingual* aspects of plurilingual competence that are covered by multiple approaches are (...) an end in itself, of which proficiency could be measured? What would be, from this point of view, a "plurilingual" competence at A2 or C1 level?"

The question that J.-F. de Pietro is asking is, in my opinion, twofold, since it concerns both the relevance:

- to make plurilingual competence the object of assessment;
- to reduce this assessment to a matching of levels using standardized descriptors.

The first part of this questioning was raised above. The other part actually refers to the evaluation design underlying the work on the evaluation of plurilingual competence. However, while the conception of language is generally not dealt with to any great extent, the conception of evaluation is not dealt with at all, even though this point is fundamental, especially in the perspective of epistemological coherence defended here.

---

<sup>9</sup> "Appropriation" refers here not to possession and mastery, but to an "existential, experiential and relational process" (Castellotti 2020: 5) integrating deviations, contingencies, mixtures (the "ghosts" of other languages, to use Py's beautiful metaphor 1992) and posing as central "the place [that language] occupies in the history and the project of the person who engages in this appropriation" (ibid.).

<sup>10</sup> But can we still speak of "competence" in this case? And isn't all language competence in fact plurilingual?

What does it mean, then, to "put diversity at the heart of evaluation"? The thesis I have been trying to defend for several years is that diversity is epistemologically evicted as soon as the assessment practices and tools are based on an imaginary of transparency and control, even if they are in several languages or if they integrate trans- and interlinguistic dimensions. However, this design of evaluation currently prevails in language didactics, including work on plurilingual competence. Many of titles of books devoted to the assessment of plurilingual competence emphasize transparency, the "quality approach" or the reliability of descriptors (cf. for example Taylor and Weir 2005, Piccardo et al. 2019). This discourse is so dominant in language didactics that it has become self-evident, making us lose sight of the fact that other conceptualizations of evaluation (could) exist, which are (would be) based on diametrically opposed principles (cf. for example Ardoino and Berger 1986 or Vial 2001; cf. also the historical work of A. Desrosières on literary origin of statistic). They may be summarized roughly as follows:

- evaluation depends most of all on the teachers' and/or learners' (and evaluators') relationship to assessment (and more generally to diversity). To take a practical example and as I have already mentioned elsewhere (Huver and Springer, 2011 ; Huver to be published), an assessment tool supposed to be very qualitative (such as the portfolio, for example) can be used in an extremely controlling way — just as, conversely, a monolingual control-type tool can be understood and invested in a different way. In other words, practices and tools are based on representations and thus fall first under a "universe", an imaginary, a way of conceiving the world. Therefore, for researchers and teachers trainers, the (reflexive) understanding of the universe in which the evaluator is involved is far more important than developing, advising or prescribing tools.
- the evaluative process is fundamentally hermeneutic rather than analytic. In analytical, objective evaluation, the quality of evaluation derives from the rigorous use of tools, from the objectivity of the evaluator and from the reproducibility of his or her judgement (see for example the way that certifying bodies define the quality of their tools). On the contrary, in a hermeneutic paradigm, the quality of assessment is based on the flexible use of rules, on the committed and sensitive presence of the evaluator and on the conflict of interpretations. More precisely, Romano (2015:473), basing on Gadamer's theory, defends the idea that the interpretation (in its hermeneutical sense) does not arise from the strict, exact, right use of a method, which is not in itself able to preserve from error. To the idea of "exact" rules that could be applied mechanically, Romano opposes "the (flexible) rules of experience", which "rather require a use that itself has experience" (ibid.: 475). The central status of experience (and reflexivity) implies to admit that there is a form of "pre-methodical" understanding that passes below, beyond and through rationalized understanding and that this "pre-methodical understanding" plays a fundamental role in the assessment process. And this applies a fortiori for plurilingual competence, which, is - when defined as below – also inherently qualitative, thus interpretative / hermeneutic.

Considering evaluation from this angle is therefore the opposite of an evaluation based on a priori designed standards and ready-made psychometrically validated scales on the one hand, and that delegitimizes the evaluators' experience on the other.

#### **4. Conclusion: from epistemological consistency to ethical vigilance**

Under what conditions can we speak of effective assessment of plurilingual competence, or, more precisely, of effective plurilingual assessment?

It is clear that developing tools that continue to evaluate languages in a juxtaposed relationship is not enough (and there is a consensus among researchers interested in this matter, including in this book). It is indeed on the basis of this observation that tools have been developed that try to take into account the inter- and trans-linguistic dimensions or the mediation issues inherent in any language contact.

But if we wish to be consistent with the changes brought about by the notion of plurilingual competence in language didactics, developing "not monolingual assessment tools" is just a superficial way of dealing with this issue. It is also necessary (and from my point of view a priority) to question



the notions of language and evaluation by considering diversity as an epistemological and (therefore) political principle:

- on which underlies the interpretative (and therefore evaluative) judgement, considering that rigour is not to be found in the tools or in their (proper) use—in short, in the method and technique—, but in the involvement of the evaluator and in the plurality of interpretations that a reflective posture must allow to make explicit and to confront;
- on conceptions of what is called "language", considering it not only as a heterogeneous social phenomenon, but also as experiences of the world, which cannot, by definition, neither be controlled nor technologized.

In my opinion, this shifts questions about what the social intervention of researchers in language didactics can mean and therefore presents ethical (and not only philosophical) issues from this point of view:

**In the field of teacher training:** it is not about deciding instead of teachers what tools they should carry out—and therefore about training them to the proper use of these tools. It is not either about developing in them a "culture of evaluation" (Cuq and Davin 2016), which would suggest that only one (homogeneous) culture would be valid and that some of the teachers would be, a contrario, "evaluation uneducated". Language assessment training in which diversity is a priority means working on the basis of the evaluators' experiences and representations, in order to work primarily on becoming aware of one's collective and individual stakes and imaginaries on language and assessment—in order to question them or, on the contrary, *to better understand why they are important*. Therefore, for the researchers, thinking about a form of responsible social intervention is less about creating tools than fostering reflexivity by teachers, based on the diversity of their positions and fundamentally respectful of this diversity.

**In the field of research:** it is common to hear or read that the provision of level reference systems or tools for assessing plurilingual competence would promote its teaching (Goullier 2017: 256). The responsibility of researchers who wish to promote this plurilingual turning point would then lie in developing and disseminating such devices. From the perspective I have tried to defend here, the responsibility of researchers is based on a need for consistency, not only between teaching and assessment practices, but also between the concepts that underlie these practices. Since plurilingual competence is rooted in the notion of diversity, as defined here, it is indeed this notion that must serve as an epistemological thread to think evaluation and language notions and, on this basis, evaluative practices.

Without this epistemological coherence, the evaluation of plurilingual competence is condemned to assess only a superficial plurilingualism. And it is by firmly and insistently placing diversity at the center of the evaluation process that we can guard against some manipulations of the tools created, since diversity introduces a form of relativity that excludes any expert positioning. Thus, given that language is defined as an experience of the world and its learning as an experience of contact and otherness, language assessment (but also language teaching/learning) cannot (epistemologically) be understood solely from a technical point of view and should not (ethically) be so. Placing diversity at the heart of the assessment of plurilingual competence is therefore above all to think of epistemological coherence as a form of ethical vigilance.

## Bibliographie

- Anderson, P., 1999, *La didactique des langues étrangères à l'épreuve du sujet*, Besançon : Presses universitaires franc-comtoises.
- Anquetil M., Derivry M. et Gohard Radenkovic A., 2017, « En finir avec le Je contraint et réifié dans l'objet PEL : pour une didactique de la biographie langagière comme processus relationnel », *TDFLE*, n° 70, *La pensée CECR*, <http://revue-tdfle.fr/les-numeros/numero-70/23-en-finir-avec-le-je-contraint-et-reifie-dans-l-objet-pel-pour-une-didactique-de-la-biographie-langagiere-comme-processus-relationnel>.
- Ardoino J. et Berger G., 1986, « L'évaluation comme interprétation », *POUR*, n° 107, pp. 120-127.

- Bachman, L. F., & Palmer, A. S., 2010, *Language assessment in practice: developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.
- Blanchet Ph., 2016, *Discriminations : combattre la glottophobie*, Paris : Textuel.
- Butera F., Buchs C. et Darnon C., 2011, *L'évaluation une menace ?* Paris : PUF.
- Castellotti V., 2020, « La réception -compréhension -traduction en didactologie-didactique des langues : quelles conceptions, pour quelles orientations? », *In-pertinences*, n° 1, <https://dynadiv.univ-tours.fr/in-pertinences/debats-et-essais-522869.kjsp?RH=149138747082>.
- Castellotti V., 2017, *Pour une didactique de l'appropriation. Diversité, compréhension, relation*, Paris : Didier.
- Castellotti V. et Moore D., 2011, « La compétence plurilingue et pluriculturelle : genèse et évolution d'une notion-concept », Blanchet P. et Chardenet P., *Guide pour la recherche en didactique des langues et des cultures. Approches contextualisées*. Paris : Editions des Archives Contemporaines, pp.241-252.
- Chardenet P., 2005, *Evaluer des compétences plurilingues et interlingues*, [https://www.academia.edu/5354831/Evaluer des comp%C3%A9tences plurilingues et interlingues 2005](https://www.academia.edu/5354831/Evaluer_des_comp%C3%A9tences_plurilingues_et_interlingues_2005).
- Cheng, L., & Fox, J., 2017, *Assessment in the Language Classroom: Teachers Supporting Student Learning*, Macmillan Education.
- Coste D., 2021, « De Rüscliikon au *Volume complémentaire* ou Du risque qu'il y a à passer sous les échelles ». In Vogt, K., & Quetz, J. (Éds.). *Der neue Begleitband zum Gemeinsamen europäischen Referenzrahmen für Sprachen*. Berlin : Peter Lang. pp. 35-45.
- Coste D., 2002, « Compétence à communiquer et compétence plurilingue », dans : *Notions en questions*, n° 6, « La notion de compétence en langue », Castellotti V. et Py B. (coord.), pp. 115-123.
- Conseil de l'Europe, 2018, *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer – Volume complémentaire avec de nouveaux descripteurs*, Strasbourg : Éditions du Conseil de l'Europe, <https://rm.coe.int/cecr-volume-complementaire-avec-de-nouveaux-descripteurs/16807875d5>.
- Cuq J.-P. et Davin F., 2016, « L'évaluation, un paramètre prépondérant en didactique du français langue étrangère et seconde », *L'évaluation à la lumière des contextes et des disciplines*, Bruxelles : De Boeck, pp. 91-110.
- Davis, J. M., Norris, J. M., Malone, M. E., & McKay, T. H. (Eds.), 2018, *Useful Assessment and Evaluation in Language Education*, Georgetown University Press.
- Debono M., 2020, *La question de la technique en didactique des langues et en sociolinguistique (francophones)*, Habilitation à diriger des recherches, Université Grenoble Alpes.
- De Pietro J.-F., 2019, « Faut-il évaluer la « compétence plurilingue » ? Quelques réflexions à partir du Cadre de référence pour les approches plurielles (CARAP) », in : Gajo L. et al., op.cit.
- Desrosières, A., 1993, *La Politique des grands nombres – Histoire de la raison statistique*, Paris : La Découverte.
- Gajo L., Luscher J.-M., Racine I. et Zay F., 2019, *Variation, plurilinguisme et évaluation en français langue étrangère*, Bern : Peter Lang.
- Goullier F., 2017, « Evaluer la compétence plurilingue et interculturelle ? », in : Beacco J.-C. et Coste D., *L'éducation plurilingue et interculturelle : la perspective du Conseil de l'Europe*, Paris : Didier.
- Hidri, S. (Ed.), 2018, *Revisiting the Assessment of Second Language Abilities: From Theory to Practice*, Springer.
- Huver E., to be published, "Favoriser l'autonomisation par l'évaluation: s'autonomiser de l'évaluation ?", Candau P., Acker P. et Toffoli D. dir., *Autonomisation et évaluation*, Peter Lang.
- Huver E., 2014, *De la subjectivité en évaluation à une didactique des langues diversitaire. Pluralité, altérité, relation, réflexivité*, Note de synthèse d'habilitation à diriger des recherches, Université F. Rabelais (Tours), <https://hal.archives-ouvertes.fr/tel-01211154v1>.
- Huver E., 2014b, « Les inaccessibles en évaluation. Impensé ? Impasse ? Ferments ? », Goï C., Huver E. et Razafimandimbimanana E. (coord.), « Inaccessibles, altérités, pluralités : trois notions pour

- questionner les langues et les cultures en éducation », *Glottopol*, n°23, [http://glottopol.univ-rouen.fr/telecharger/numero\\_23/gpl23\\_03huver.pdf](http://glottopol.univ-rouen.fr/telecharger/numero_23/gpl23_03huver.pdf).
- Huver E., 2010, « J'ai pas pu corriger parce que c'est le mélange des langues » - L'évaluation de la compétence plurilingue entre résistances et tensions », *Cahiers de l'ACEDLE*, n° 8, [http://acedle.org/IMG/pdf/Huver\\_Cahiers-Acedle\\_7-2.pdf](http://acedle.org/IMG/pdf/Huver_Cahiers-Acedle_7-2.pdf).
- Huver E. et Macaire D., 2021, « Didactique de langue, didactique des langues, didactique du plurilinguisme : évolutions, enjeux, questions », *Recherches en didactique des langues et des cultures*, n° 18-2, <https://journals.openedition.org/rdlc/9673>.
- Huver E. et Springer C., 2011, *L'évaluation en langues*, Paris : Didier.
- Kunnan A.-J. (dir.), 2014, *The Companion to language assessment*, John Wiley and sons ed.
- Lenz P., 2006, *Überlegungen zur Sprachkompetenzbeschreibung und Testvalidierung im Projekt HarmoS/Fremdsprachen*, [https://www.researchgate.net/publication/42384750\\_Uberlegungen\\_zur\\_Sprachkompetenzbeschreibung\\_und\\_Testvalidierung\\_im\\_Projekt\\_HarmoS\\_Fremdsprachen](https://www.researchgate.net/publication/42384750_Uberlegungen_zur_Sprachkompetenzbeschreibung_und_Testvalidierung_im_Projekt_HarmoS_Fremdsprachen).
- Lenz P., & Berthele R., 2010, *Assessment in Plurilingual and Intercultural Education*, Conseil de l'Europe, Strasbourg, [http://www.coe.int/t/dg4/linguistic/ListDocs\\_Geneva2010.asp#TopOfPage](http://www.coe.int/t/dg4/linguistic/ListDocs_Geneva2010.asp#TopOfPage).
- Madaus, G. F., 1990, *Testing as a Social Technology*, Boston : Boston College.
- Maurer B. et Puren C., 2019, *CECR : par ici la sortie !*, Paris : Éditions des archives contemporaines, <https://www.christianpuren.com/mes-travaux/2019d>.
- Mc Namara T. et Roever C., 2006, *Language Testing: The Social Dimension*, Oxford (UK) : Blackwell Publishing.
- Medioni M.-L., 2016, *L'évaluation formative au cœur du processus d'apprentissage*, Lyon : Chronique sociale.
- Ollivier Ch., 2019, « Évaluer les compétences en communication plurilingue – le cas de l'intercompréhension », *Les Langues Modernes* n° 4, « Évaluation et certification en langues : tensions, évolutions, perspectives ».
- Piccardo, E., North, B., & Maldina, E., 2019, "Promoting innovation and reform in language education through a Quality Assurance template for CEFR implementation", *Canadian Journal of Applied Linguistics*.
- Prieur, J.-M. et Volle, R.-M., 2016, « Le CECR : une technologie politique de l'enseignement des langues », *Educations et sociétés plurilingues*, n° 41, pp. 75-87.
- Py B., 1992, « Acquisition d'une langue étrangère et altérité », *Cahiers de l'ILSL*, n° 2, pp. 113-126, <http://www.unil.ch/webdav/site/clsl/shared/CILSL02.pdf>.
- Romano C., 2015, « La règle souple de l'herméneute », *Critique*, n° 817-818, pp. 464-479, <http://www.cairn.info/revue-critique-2015-6-page-464.htm>.
- Romano C., 2010, *Au cœur de la raison : la phénoménologie*, Paris : Gallimard, Folio.
- Shohamy E., Or, I. G., & May, S. (Éds.), 2017, *Language Testing and Assessment*, Springer.
- Shohamy E., 2001, *The power of tests. A critical perspective on the uses of language tests*. Essex : Pearson Education.
- Soubre V., 2021, *L'agir évaluatif partagé : une ressource dynamique pour l'enseignement et l'apprentissage des langues étrangères dans un contexte d'enseignement supérieur*, thèse de doctorat, université de Genève, <https://archive-ouverte.unige.ch/unige:150813>.
- Spolsky B., 1995, *Measured word. The development of objective language testing*. Oxford: Oxford University Press.
- Taguchi, N., & Kim, Y. (Eds.), 2018, *Task-Based Approaches to Teaching and Assessing Pragmatics*, John Benjamins Publishing Company.
- Taylor L. et Weir C.-J. (Eds.), 2005, *Multilingualism and Assessment: Achieving Transparency, Assuring Quality, Sustaining Diversity—Proceedings of the ALTE Berlin Conference*, May 2005.
- Vial M., 2001, *Se former pour évaluer, Se donner une problématique et élaborer des concepts*, Bruxelles : De Boeck.