Assessment in Plurilingual and Intercultural Education

Peter Lenz and Raphael Berthele

Document prepared for the Policy Forum 'The right of learners to quality and equity in education – The role of linguistic and intercultural competences'

Geneva, Switzerland, 2-4 November 2010

LIST OF SATELLITE STUDIES ACCOMPANYING THE

Guide for the development and implementation of curricula for plurilingual and intercultural education

1. *A curriculum perspective on plurilingual education*
   Jan van den Akker, Daniela Fasoglio, Hetty Mulder

2. *Assessment in Plurilingual and Intercultural Education*
   Peter Lenz & Raphael Berthele

Table of contents

Outline

This satellite to the *Guide for the Development and Implementation of Curricula for Plurilingual and Intercultural Education* focuses on assessment aspects. First, a brief overview is given of the various aims of assessment in plurilingual and intercultural education. The authors' understanding of assessment is then briefly outlined. Finally, the assessment of selected competences having to do with core aspects of plurilingual and intercultural education is presented and discussed.

1 What are the aims of assessment in plurilingual and intercultural education?

The introduction of plurilingual and intercultural education is a whole-school project extending over some time, which aspires to further improve the quality of education and to have a positive impact on society at large (cf. "Guide", section 1.3.1. and Beacco/Byram 2007). The language subjects (language of schooling and foreign languages) are most directly concerned; here, it is intended to use potential synergies and to target and combine the efforts made in language and (inter-)cultural education more efficiently and effectively. Non language-specific subjects having to do with language and (inter-)cultural knowledge and skills – i. e., in principle (almost) all school subjects – are also to be included in this overarching approach. Assessment is an integral part of individual learning and of the education system. Finding suitable ways of assessing plurilingual and intercultural competence is a step towards its integration in educational practice. Depending on the degree of implementation of plurilingual and intercultural education, and the curricular options chosen, a greater or lesser number of people and existing structures are concerned, and the need for specific innovations and changes arises to a greater or lesser extent. Assessment needs change accordingly. They may relate to rather different aspects of the renewed system. Institutional and organizational points of view are not considered in the following overview which focuses instead on aspects closely related to the objectives of plurilingual and intercultural learning as they appear in the Guide.

A **Assessment needs regarding the over-arching goals of plurilingual and intercultural education:** are the learners' competences in a range of languages generally as good as expected? Are the learners competent as learners of languages and (inter-)culture, as well as self-directed (autonomous) to the desired degree? Are the various stakeholders aware of regional and minority languages and the language varieties surrounding them, and do they appreciate them (cf. Council of Europe (2009)? Are they open towards and familiar with different cultures, including their own? Do they generally wish to intensify their knowledge of and relationships with 'others'? Are minorities well integrated in the school community? Are socially less favoured students becoming more successful? Do the curricula ensure the right to plurilingual and intercultural education (Coste et al, 2009)? Finally, an important criterion from an economical point of view: are integrated curricula that emphasise plurilingual and intercultural approaches for teaching and learning language and culture more effective and efficient than conventional arrangements?

Remark: organizational and economical aspects, as well as the monitoring of educational systems and other large-scale and/or accountability-oriented assessment purposes are not discussed further in this document which focuses on aspects of assessment that concern the development of individual learners.

B **Assessment needs regarding competences that help learners to participate successfully in social life** (cf. section 1.2.1 in the *Guide*)**:**

1) Ability to use (and thereby improve) their plurilingual repertoire:

- language use for educational purposes: use of languages, varieties and discourse genres according to the requirements defined by different school subjects (language and others); this encompasses the language(s) of schooling as well as other languages, and, in particular, language varieties for (subject-)specific purposes;

- language knowledge for uses beyond school: contacts with users of other languages in the home environment or elsewhere; use in prospective vocational or professional training.

Plurilingual education in a broad sense relates to language competences of all kinds, including language competence in a single language. Plurilingual education in a narrower sense emphasizes the ability to mobilize the language repertoire as a whole, to use existing

competences transversally, i. e. to recombine existing knowledge and skills in any language(s) in order to respond flexibly to needs that arise in a multilingual environment. Objectives regarding plurilingual competence (in a narrower sense) may have to do with:

- communicating orally in plurilingual settings, e. g. engage in polyglot dialogue; use code-switching and code-mixing as communicatively and contextually functional devices;

- using multiple input sources in different languages in order to accomplish productive or interactive tasks in a strong language;

- exploiting a *profile* of unevenly developed language skills across several languages which, as a whole, suits the communicative needs, e. g. excellent overall skills in language A; strong receptive skills in language B; strong oral skills (receptive and productive) in language (or variety) C, good reading skills in language D; potential reading skills in several closely related languages that s/he never actually studied;

- mediating between languages, e. g. translating and interpreting; passing on in simple words in language B the content of a text read in language C;

- using available knowledge of all kinds from previous language learning in order to understand texts in genetically related languages (intercomprehension, e. g. between Romance, Slavic, Germanic languages).

Although multilingual contexts are a reality and the simultaneous use of more than one language and/or variety has distinct advantages in many settings, code-mixing and code-switching are usually considered as non-authorized forms of language use in schools. We do not deny that monolingual regimes in language teaching can be appropriate at certain times and depending on the learning outcomes defined by the curricula. However, we argue for a plurality of mono- and plurilingual regimes that also takes into account the simultaneous use of the plurilingual repertoires of the participants involved. In this respect, fixed mindsets need to be replaced by more realistic views of actual language use and language norms.

2) Ability to use (and thereby enlarge) their pluricultural and intercultural repertoire in interaction with otherness:

- ability to identify with and have a share in different cultures (pluricultural competence);

- ability to question the seemingly self-evident brought about by one's own cultural background;

- ability to react in non-ego, ethno or socio-centric ways to other cultures or sub-cultures;

- ability to analyse cultural otherness;

- ability to act as a mediator in intercultural encounters, etc.

It is obvious that languages cannot be used successfully in authentic contexts without at least some intercultural competence. In this view of intercultural competence as an ancillary competence, it is uncontested that language education must include elements of intercultural education. But in the Guide, intercultural competence is considered as a competence area in its own right.

Intercultural competence, at its core, has to do with the integration of 'otherness' in one's thinking and actions. In this respect, it is significantly different from a concept of *pluricultural* competence, which highlights the plurality of cultures one may identify and is familiar with. Developing intercultural competence means building an intercultural identity to some degree. Intercultural competence is not necessarily linked to the knowledge of other languages but knowledge of languages considerably widens the horizon of possible intercultural experience and interpersonal contact, which may lead to a different quality of intercultural competence. Knowledge of a *lingua franca* can facilitate intercultural experience a great deal. Intercultural experience does not start 'somewhere abroad'; it starts by acknowledging and paying attention to the unfamiliar. There is ample opportunity for intercultural experience in today's far-from-unitary cultures that prevail in many parts of the world. We consider plurality as the default case in human cultures, since variation, contact and dynamics of languages and cultures are – diachronically and synchronically – ubiquitous and the idea of homogenous linguistic and cultural spaces has been particularly influential in European nation-building processes of the

18th and 19th centuries[1]. It is thus important to realize that cultural differences do not necessarily interlock with language borders.

C **Assessment needs regarding competences and attitudes that help learners to develop as critical intercultural individuals and citizens**

- openness and willingness to experience cultural and linguistic otherness;
- ability to reflect about intercultural competence and its uses;
- acknowledging the value of cultural and linguistic diversity, variation and otherness;
- acknowledging participation in a multicultural and multilingual society as a right and responsibility at the same time, etc.

In Byram's concept of *Education for Intercultural Citizenship* they are taken further where the mission of foreign-language teaching is extended through the addition of (critical, democratic) 'values' and by making it part of education for democratic citizenship (Byram, 2008: 227-8). According to this concept, (foreign) language teaching

> brings a specific additional contribution to education for (democratic) citizenship. That contribution is captured in the term 'intercultural citizenship', a focus of citizenship education on the understanding and action involved when one is a member of an international society, especially of an international civil society. (Byram, 2008: 229)

D **Assessment needs regarding competences that help learners to take responsibility and develop as language users and learners**

1) Understanding and awareness of languages:
   - ability to reflect on communication and human language;
   - understanding the way (different) languages work;
   - ability to analyse linguistic data;
   - awareness of the historical embedding, the diversity, variability, and the creative potential of languages;
   - understanding how discourse contributes to the construction of information, opinions, ideas, ideologies and consciousness.

These aspects have been elaborated most notably in projects concerned with the *Awakening to languages* (e. g. EvLang; Janua Linguarum)*.* This approach intends to open up the world of (interacting) cultures, languages and language learning mainly to novice learners. Within this concept, developing usable communication skills is seen as a second-level goal. *Awakening to languages* is concerned with first steps into interculturalism, general language education, and establishing a basis for effective learning of any language.

2) Ability to learn languages; ability for self-directed (autonomous) and life-long learning:
   - being experienced in language learning and the use of a variety of languages for different purposes, including ludic and aesthetic activities, etc.
   - acknowledging the value of all language competence, even partial;
   - awareness of one's own existing language competences;
   - awareness of and ability to use language learning strategies;
   - ability to self-assess language competence;
   - ability to see potential advances and to plan further learning;
   - awareness of possible transfers of knowledge and skills to other contexts, etc.

---

[1] Although undoubtedly the ideology has emerged earlier and had its impact in Eurpean-based colonization processes outside Europe as well.

The development of learners as such has been a long-standing objective in Council of Europe work and publications, and not only in the context of language teaching and learning. It is an area of competence in which many synergies can be expected when the teaching of all subjects in school is integrated and coordinated. Many of the competences that make up the ability to learn can be expected to prove highly transferable from one learning context to the other.

E Assessment needs regarding competences that help learners to acquire new knowledge in their studies in plurilingual settings

- Ability to use (and thereby diversify) their plurilingual, pluricultural and intercultural repertoire in knowledge building :
- ability to reuse (and not simply reformulate) in one language subject knowledge acquired in other languages ;
- ability to carry out the tasks and cognitive operations required by one or several disciplines in more than one language while respecting the disciplinary conventions linked to these languages;
- ability to use oral and/or written sources in several languages in order to build new knowledge in one or several disciplines;
- ability to use the epistemological approach to a given subject, which may vary according to educational cultures, in order to acquire a richer and more diverse knowledge of the subject;
- ability to build more solid knowledge through bi- or plurilingual teaching, also through the linguistic opacity resulting from the use of 2nd or 3rd languages in the various subjects. Indeed, the absence of linguistic transparency can also contribute to building subject knowledge, as it leads the learner to think about the meaning of the words, the concepts behind those words and their exact definition;
- ability to identify the cultural point of view particularly important in certain subjects (cf. history)
- ability to summarise different cultural points of view on a given subject.

The preceding paragraphs dealt with "language as a subject" dimension. A second dimension – "language in other subjects" – is relevant to the increasingly frequent situations where several languages of schooling coexist. This type of reflection may concern:

- education systems, schools or even classes where bilingual teaching is present, whether in contexts where a regional, minority or migration language is used for subject learning in parallel with a majority school language, or situations where a foreign language is used in similar fashion (for example in CLIL/EMILE contexts)
- universities where classes are given in more than one language.

The "language in other subjects" dimension, whose (pluri)linguistic, (inter)cultural and cognitive implications are very complex, needs substantial development, including in terms of research. For this reason this study only refers to it as an important component of plurilingual and intercultural education and makes no attempt to take account of it in its consideration of plurilingual and intercultural competences.

Operationalisation

The above enumeration of aspects of plurilingual and intercultural education that might be worth assessing needs to be further operationalised before assessment can actually be envisaged and suitable forms of assessment can be allocated.

The aspects mentioned in A through D can be operationalised more easily and with more confidence whenever specific *reference frameworks* exist to which they can be related. This is the case for some aspects more than for others.

The *Common European Framework of Reference for Languages* (CEFR, Council of Europe, 2001) lends itself as a framework mainly for the communicative use of foreign languages (section B 1, above), most notably one foreign language at a time. For this most familiar case, language activities and corresponding competence resources are described and attributed to a scale ranging from A1 to C2.

Conceptually, language use involving more than one language forms part of the classification provided insofar as mediation[2] is one of the four main communicative modes (in addition to reception, production and interaction) the language activities are attributed to. However, mediation is that part of this classification that was not illustrated or scaled. Other forms of plurilingual language use do not appear in the illustrative scales, either. Nevertheless, the CEFR provides a competence model and a descriptive system that is useful even beyond its present scope. The action-oriented approach combines basic descriptive categories like task, text, topic; strategy; language processes, general competences and communicative language competences[3] that may be applied to language activities not yet described or scaled. This may concern activities having to do with language use for specific purposes (e. g. the first language of schooling or other languages used in so-called "non-language subjects") or language use involving more than one language for educational purposes or for uses outside the school context.

Descriptive and analytical systems such as that reflected in the CEFR's action-oriented approach are particularly useful for teaching and assessment. In teaching (and learning) they form a basis for breaking down learning objectives into more manageable pieces like lexical units, pragmatic aspects, phonetics, characteristics of discourse genres, the necessary knowledge about the world etc. Classroom-based assessment may relate directly to micro-objectives derived from such smaller teaching/learning units (vocabulary tests etc.), or else, when performances on complex tasks are assessed, criteria corresponding to these particular aspects may be applied (cf. Table 3 of the CEFR or other assessment grids).

As concerns intercultural competence (sections B 2 and C, above), more or less elaborated approaches to reference frameworks exist but empirically founded scales of intercultural competence are yet to be developed. Chapter 5.1 of the CEFR (*General competences*) provides some detail on intercultural competence in so far as it is relevant for successful communication. Different types of intercultural competences are attributed to the familiar *savoir* categories: *Savoir* (declarative knowledge, e. g. cultural knowledge; knowledge about cultural differences); *Savoir faire* (skills and know-how, e. g. acting according to the rules; getting in contact); *Savoir être* (existential competence, e. g. attitudes like openness; willingness to relativise).

Byram (1997) presents a more general framework of intercultural competence: the concept of *Intercultural Communicative Competence* (ICC). The understanding of intercultural competence put forward in the Guide, corresponds very much to ICC. This concept expresses its objectives in similar categories (i. e. different types of cognitive resources) as the CEFR, interpreting and extending them in important ways:

- attitudes (*savoir être*): curiosity and openness, readiness to suspend disbelief about other cultures and belief about one's own;

- knowledge (*savoirs*): of social groups and their products and practices in one's own and in one's interlocutor's country, and of the general processes of societal and individual interaction;

- skills of interpreting and relating (*savoir comprendre*): ability to interpret a document or event from another culture, to explain it and relate it to documents from one's own;

- skills of discovery and interaction (*savoir apprendre/faire*): ability to acquire new knowledge of a culture and cultural practices and the ability to operate knowledge, attitudes and skills under the constraints of real-time communication and interaction;

- critical cultural awareness/political education (*savoir s'engager*): ability to evaluate critically and on the basis of explicit criteria perspectives, practices and products in one's own and other cultures and

---

[2] The CEFR introduces mediation as follows: 'In both the receptive and productive modes, the written and/or oral activities of mediation make communication possible between persons who are unable, for whatever reason, to communicate with each other directly. Translation or interpretation, a paraphrase, summary or record, provides for a third party a (re)formulation of a source text to which this third party does not have direct access.' (Council of Europe, 2001:14)

[3] 'Competence' is not used in the same way in the CEFR as in the Guide. The CEFR's 'general competences' and 'communicative language competences' would be considered as 'resources' in the Guide; 'language competence' is used in the Guide as 'ability to use language', which in the CEFR would be called 'language proficiency'. In general English, 'competence' may be used in either way or as a more general term that subsumes both. In this document 'competence', 'ability', 'skill' and 'proficiency' are all used to refer to 'ability to use language' unless indicated otherwise.

countries; as a result of such evaluations, the learner may engage in interactions and negotiations and act as a mediator. (Byram, 1997: 47-54; Byram, 2008: 230-33)

Meanwhile, Byram and colleagues have presented attempts to transform these aspects of ICC into concrete tasks and impulses for reflection (INCA project[4]; Autobiography of Intercultural Encounters[5]). Moreover, Byram and his INCA project partners have described six constituents of intercultural competence, combined in pairs to form three strands, and defined in three (tentative) levels. These strands include:

- 1st strand: *openness:*
  respect for otherness (ability to look at all customs and values from a distance, regarding them at the same time as worthwhile in their own right),
  tolerance of ambiguity (ability to accept ambiguity and lack of clarity and deal with it constructively);

- 2nd strand: *knowledge*:
  knowledge discovery (ability to acquire and actually use cultural knowledge),
  empathy (ability to intuitively understand what other people think and how they feel);

- 3rd strand: *adaptability:*
  behavioural flexibility (ability to adapt one's own behaviour to different requirements and situations),
  communicative awareness (ability to identify and consciously work with communicative conventions).

The three levels indicate to what extent the six constituents are developed and put into action:

- basic competence: individual is reasonably tolerant, willing to interact successfully but responds only to events instead of planning for them;

- intermediate competence: individual prefers responding in a neutral way to difference, and is prepared to respond and adapt to the demands of unfamiliar situations;

- full competence: individual is confident enough of his/her own position in order to take a polite stand over issues and is constantly ready to use a large repertoire of strategies, knowledge and skills to deal with difference.

There are attempts to further dissect and structure intercultural competence which may eventually help to identify finely grained and scaled operationalised objectives. The CARAP project, still work in progress, is among the most notable of these attempts (cf. also Beacco 2004).

Language awareness (section D 1) has been topical since the 1980s. An abundance of materials was produced in the context of the EvLang project hosted by the European Centre for Modern Languages (ECML) in Graz. The *Janua Linguarum* report contains an extensive bibliography indicating many sources. Language Awareness consists of various concepts, some of which have been broken down into classroom activities, but an agreed reference framework, preferably displaying progression or levels, does not exist. Moreover, firm links between these awareness activities and language learning for plurilingual communication have yet to be established.

As concerns the ability to (consciously) learn languages (section D 2), concepts exist in addition to systematic approaches to some of the aspects. Also, concrete examples and materials have been provided in textbooks and publications focusing on the aspects concerned. However, a comprehensive, agreed or even empirically grounded framework does not exist.

Classical texts on the theory and practice of learner autonomy that continue to provide relevant reading include, for example, Holec (1981/79), Dam (1995) or Little (1991); Little and Dam are both actively contributing at the present time. The ability to self-assess is a prerequisite to learner autonomy insofar as the abilities to assess progress and plan further learning accordingly are fundamental components of the ability to take learning in one's own hands – which characterises learner autonomy. The following

---

[4] INCA is a Leonardo da Vinci project that intends to make available tools for the assessment of the intercultural competence of adults in an intercultural work environment. It encompasses introductory texts and manuals; a framework based on theory; test instruments – based on the interpretation of intercultural scenarios, and on role-play activities – as well as a portfolio for self-reflection, assessment and documentation. Some materials can be used not only on paper but also on-line.

[5] The *Autobiography of Intercultural Encounters* was developed on behalf of the Council of Europe as an ICC-specific addition to the *European Language Portfolio*.

texts are helpful in gaining an overview of the field of self-assessment: Oskarsson (1984), Dickinson (1992), Oscarson (1997), Little (2005).

In the field of language learning strategies, again, 'classics' are useful points of departure: they provide an overview, help structure the field and may help to identify the objectives of various approaches to grading - see Oxford (1990), O'Malley & Chamot (1990). However, they are neither curricula nor reference frameworks featuring dimensions and levels.

When defining objectives and assessing learning strategies, it is necessary to keep the complex nature of strategies in mind: on the one hand, having a repertoire of strategies does not necessarily mean that an individual is able to consciously reflect on or communicate about all these strategies (as they are only potentially conscious); on the other hand, strategies are known to be differentially effective – 'one suits all' does not apply in this context. Furthermore, using strategies is not intrinsically good, as cases of strategy overusers show clearly. Consequently, appropriate care has to be taken in the assessment of learning strategies and their use by learners.

The above implies that aspects of plurilingual and intercultural education that lend themselves to assessment can be operationalised at varying degrees of refinement. In a school context, for example, an operationalisation may correspond to an objective of a task or single teaching unit, an area of knowledge or skill developed over some time, or the educational outcome after a few years. As assessment comes into play, it is important to be clear about what it is that is assessed. This is where the notion of construct comes into play. – Bachman & Palmer (2010:43) define *construct* as follows:

> For our purposes, we can consider a construct to be the specific definition of an ability that provides the basis for a given assessment or assessment task and for interpreting scores derived from this task. The construct definition for a particular assessment situation becomes the basis for the kinds of interpretations we can make from the assessment performance. In designing, developing, and using language assessments, we can define the construct from a number of perspectives, including everything from the content of a particular part of a language course to a needs analysis of the components of language ability that may be required to perform language use tasks in a target language use domain, to a theoretical model of language ability.

Bachman and Palmer developed their definition with the assessment of language ability in mind. However, we argue here that it is appropriate and possible to extend its field of application to the clearly wider area of plurilingual and intercultural education as far as abilities/competences (including attitudes) are concerned. In all these cases, conclusions regarding an individual's competence on the basis of observable behaviour (in assessments) are justified only if the link between the two is made transparent.

2 Assessment

While the various potential objects of assessment in plurilingual and intercultural assessment have been the topic of the first chapter, the following sub-chapters will now discuss different aspects of assessment: fundamental types, quality aspects, purposes, assessors as well as assessment instruments.

2.1 What is assessment?

Competences, awareness and attitudes are not usually observed directly. Appropriate instruments and methods have to be used in order to open up ways of gaining insight into the various constructs. This is no trivial task. If assessment is understood in a broad sense and not limited to testing, ways can be found to assess even those phenomena which may not seem to lend themselves to assessment very easily. Whatever the assessment method is, it should be born in mind that all of them have their specific limitations.

In the field of *language* assessment (which, again, is our point of departure), the use of terminology is far from consistent, even among experts. In their authoritative book *Language Assessment in Practice,* Bachman and Palmer (2010) conceptualize assessment as collecting information in order to help people make decisions. In this very broad sense, *assessment* subsumes other related terms, such as

evaluation, *test, alternative assessment* and the like. Also, *assessments* (plural) is used to denominate single instances of assessment, such as an oral interview or filling in a self-assessment checklist. – Bachman and Palmer's use of *assessment* and *assessments* is adopted in this document.

Bachman & Palmer (2010:28ff.) make a fundamental distinction between implicit and explicit assessment. *Implicit* assessment usually 'just happens'. It often accompanies one's own practice (e. g. when presenting a topic), or somebody else's practice (providing instantaneous corrective feedback), and adapts dynamically to ever new situations. *Explicit* assessment, on the contrary, is labelled as such for all those involved; in a classroom context, it is separate from teaching. In self-directed learning it may involve the use of a questionnaire, a checklist or other suitable assessment instruments. Implicit and explicit assessments (i. e. instances of assessment) often serve different purposes and are linked to different kinds of decision; typically, they also entail consequences that differ regarding their curricular and biographic impact. While implicit assessment of one's own presentation, for example, may simply cause one as a presenter to speak louder, explicit assessment of the same presentation by experts may have as a consequence that a presenter is successful in passing a course of study and can go on to another institution. In the remainder of the document, the focus will be mainly on explicit assessment.

2.2   Demands on assessments

It is important for assessors and users of assessment to adhere to the principle that decisions having serious consequences are based on assessments that form a suitable and sufficiently firm basis for such decisions. Accordingly, much less rigor is called for when the consequences drawn from an assessment have little potential to do harm to any of the stakeholders.

Bachman & Palmer (2010:30) suggest that in all (explicit) assessments an *assessment justification* should be brought forward, consisting of an *Assessment Use Argument* (AUA) and *supporting evidence*. The AUA makes clear how an assessee's performance, the score or result on the assessment, the interpretation regarding the ability in question, the decisions made on that basis and the consequences that may follow interrelate. The evidence collected during test development, use and analysis serves to support the claims made in the AUA.

Justifications of test use and interpretation have often been called for in recent years, most strongly by sceptics of migrant language testing (testing their knowledge of the local language – a high-stakes situation) and the critical language testing movement[6].

Messick's innovative work featuring a comprehensive view of validity (Messick, 1988), including ethical viewpoints, has greatly influenced the present developments in educational testing, as has   also Bachman and Palmer's approach that has been developing since the 1990s. Bachman & Palmer (2010:31) formulate four 'fundamental principles' for (explicit) assessments, three of which relate to *assessment justification*, the fourth to collaboration and participation:

1 The need for developers and users of language assessments to be able to *justify* to stakeholders the uses (decisions, consequences) that are made of assessments,

2 The need for *a clearly articulated and coherent Assessment Use Argument* (AUA), linking assessment performance to interpretations and to intended uses,

3 The *provision of evidence* to support the statements in the AUA,

4 The need for *collaboration among all stakeholders* during the process of assessment development and use.

These principles form a very valuable basis for an assessment culture that is marked by transparency, coherence, responsibility, rational argument, use of evidence, collaboration and dedication to good and fair practice. As implied in the first paragraph of this sub-chapter, adhering to these principles does not at all mean that any assessment needs to be highly standardized and justified in written form in order to be acceptable. It is rather *proportionality* that is asked for: any assessment has to account for the principles in some form. While the requirements for assessments that have no serious social consequences may be rather relaxed, they are very strict in cases where life or career decisions are at stake. Moreover, the impact of assessments without immediate serious consequences should not be

---

[6] See for example Shohamy (2001) or McNamara & Roever (2006).

underestimated, since they undoubtedly contribute to expectation effects (Rosenthal and Jacobson 1992) and self-fulfilling prophecies which still are a reality in educational systems. In any case, the application of the aforementioned principles can be expected to have a positive impact on the quality of assessment practice as they contain the kind of questions and criteria that drive quality assurance.

2.3 Why use assessments?

Users of assessments may differ quite fundamentally in their motives to engage in assessments. They may want to find out whether teaching has been effective, to motivate learners to study harder, to intimidate them even, to issue diplomas to graduates etc. The authors of this study are aware that the quality of actual assessments often does not meet very high standards, or that results are abused for purposes they had not been designed for. In social practice, all forms of assessments bear the potential to serve the (hidden) agenda to document differences in order to legitimate social inequalities. Nevertheless, we believe that there is still sufficient reason to trust in a fundamentally ethical motivation for assessment. We thus share to a large degree the optimistic position postulated by Bachman and Palmer (2010:26): 'The primary purpose of all assessments is to collect information to help people make decisions that will lead to beneficial consequences.' It must be underlined, however, that this kind of optimism is justified only if test users adhere to principles such as those presented above. It is quite obvious that the 'beneficial consequences' mentioned in the citation cannot be taken to mean that an assessment will make everyone happy but rather that assessments provide the information needed to make the kind of decisions that will have the most beneficial consequences overall, considering all stakeholders[7].

As mentioned before, the consequences drawn from decisions based on assessments can be anything from serious to harmless. So, it usually makes a considerable difference whether an assessment is basically *formative* or basically *summative* in nature.

Formative assessments are part of a learning cycle: collecting information is followed by an evaluation of the compiled evidence and subsequent decisions leading to consequences (modifications etc.) on learning and/or teaching.

Summative assessments gather information in order to make well-informed concluding decisions like pass-fail on a course taken (often based on an achievement test) or the attainment of a particular proficiency level (often based on a proficiency test). Summative assessments cannot be strictly separated from formative assessments if they are used to guide further teaching and/or learning. Whenever summative assessments are carried out in order to *certify* that a learner knows or is able to do specific things, care has to be taken that the competences concerned are actually valued within the respective social context. For example, it should be ascertained that a document stating some degree of skill in listening and speaking but none in reading and writing is understood as a record of achievement, not of deficits. The same is true, for example, for a proficiency descriptor such as the following: 'oral interaction about everyday things is quite possible if a learner can code-switch between target language A and his or her L1'; the ability to communicate (only) thanks to code-switching is still perceived by many as a sign of competence rather than incompetence. And also, in settings where heritage language teaching is offered in the curriculum and is also assessed, the documentation of heritage language skills, although fundamentally in line with our general idea of plurilingual competence, can be a socially stigmatizing practice, if there is no widespread consensus amongst the gatekeepers of education and employment that skills in these languages are valuable resources. Making available assessments that relate to plurilingual competence can help promote plurilingual competence and practice but care has to be taken that time and opportunity are allowed for conventional mindsets to change. We therefore call for a careful, informed application of assessments and their certification. Otherwise the intended 'beneficial consequences' will not ensue despite good intentions and good work.

2.4 How can learners be assessed?

Assessments are performed in specific constellations of assessees, assessors and assessment instruments. These performances, again, are inserted in social settings.

---

[7] For example, a fail decision on an examination may not help a candidate's career but ensure that the examination concerned will remain respected as evidence of high ability for all those who have passed.

The nature of the *assessment instruments* may vary greatly, also in terms of their degree of standardization or formality. The following are some examples of assessment instruments arranged in order of increasing degree of standardization:

- a flexible set of metacognitive strategies used to consciously reflect on success in a communicative task, an intercultural experience, a strategy chosen to learn vocabulary etc.;

- a questionnaire supporting reflection and self-assessment of intercultural encounters;

- a writing task and a corresponding set of discourse type-specific criteria for self and peer-assessment relating to relevant aspects of communicative performance (e. g. an oral presentation; a formal letter);

- a generic observation grid of criteria for teacher assessment of oral interaction in the presence of two different languages;

- a list of scaled descriptors of listening comprehension ability at adjacent levels for use in collaborative learner-teacher assessment;

- validated test tasks, complete tests and scoring rubrics produced according to agreed specifications, etc.

Expert assessment based on standardized instruments is required whenever high-stakes consequences follow from the decisions taken on the basis of the information gathered through the assessment. In such assessments validity and the demonstration of validity are crucial. A high degree of validity in test development and use can be reached and demonstrated, following Bachman & Palmer (2010), by building an Assessment Use Argument and providing supporting evidence emerging from the development process and research accompanying test use. If less is at stake, the requirements on the assessment instruments are relaxed but never completely abandoned.

Bachman & Palmer's useful approach treats assessment instruments as integrative parts of the assessments they are part of and which serve (beneficial) social functions in their respective settings. Assessment instruments are not *a priori* good, beneficial, or the opposite, but suitable or not with regard to the decisions to be taken and the consequences that follow from them.

Figure 2.1   Inferential links guiding assessment development and justifying assessment use: from Consequences to Performance on assessment tasks and back (Bachman & Palmer, 2010, p.91, fig. 5.1)
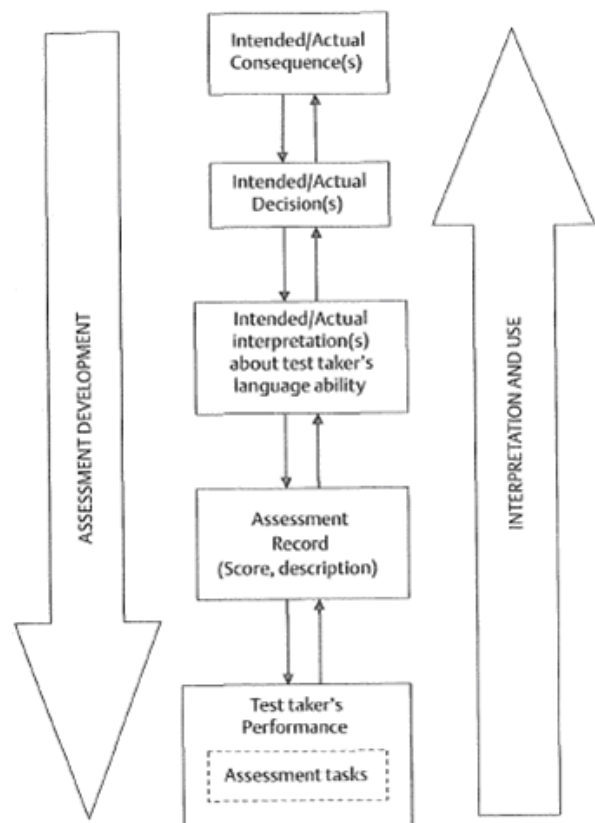
The development of an assessment takes the intended decision(s) and their real-world consequences as a point of departure for the development work and in the Assessment Use Argument (AUA) accompanying this work. For example, such a decision may be: understands introductory lectures in Spanish and English in his/her field of study well enough to benefit from them content-wise – yes/no. Depending on the answer, a counsellor at a language centre encourages a student to spend a semester abroad or to study the languages some more before leaving (consequences). In a next step, two language teachers discuss and state what kind and how much evidence they need in order to make a sufficiently reliable prediction regarding the student's real-world language skills. Since they have only limited experience with this, they decide to collect evidence and keep a record of the success of departing students. Once the assessment developers know what their assessments are supposed to yield, they draft detailed test specifications and start to design and try out with learners concrete assessment tasks (relevant for listening to lectures) and correction keys, and define pass/fail or other standards. Whenever candidates are assessed, the information thus gathered is taken into account for validation purposes and as evidence supporting the AUA, etc. Assessments developed and documented roughly according to the steps outlined and illustrated in this paragraph can be used and interpreted in order to take decisions with some confidence, depending on the degree of rigor applied[8].

The role of *assessors* may be assumed by the learners themselves (self-assessment), peers (peer assessment), teachers/facilitators (teacher assessment) or specialized assessors (expert assessment). Any combination of these possible assessors may engage in collaborative assessment. Assessors may function as assessment(-task) providers, interlocutors and/or raters.

- Self-assessment is suitable for formative assessment as well as summative assessments, provided they entail no negative consequences for the learner. Training learners to assess themselves is often considered an objective in itself since it is an integral part of learner autonomy and prepares for life-long learning. Learners, especially learners having experience with self-assessment, are often capable of insights that are not accessible to others. Only they can provide certain information on what they are able to do in real-life contact with other languages and cultural diversity, how open and interested they are to get in contact, what problems arise, what they feel, etc. Also, in a classroom context, they are in a privileged position as assessors in many respects: they know what topics and teaching sequences motivate them, experience how helpful different learning strategies are to them, can indicate immediate needs, etc. Self-assessment is known to work best when learners can relate concrete assessments back to actual experience, and when they fully understand and are familiar with the criteria they are expected to apply. Also, as mentioned above, the results of self-assessments normally should not lead to any kind of negative sanctions since learners cannot be expected to disadvantage themselves. However, it is conceivable that in collaborative classrooms with a well-rooted feedback culture, learners' self-assessments can even contribute to assessments that may have serious consequences.

- Peer assessments should be used for formative assessment only. Peer assessment serves a double purpose: 1) in addition to their own judgments, learners receive feedback from someone with an outside view who is familiar with the same tasks and criteria; peers can concentrate on one person only, teachers often have to deal with large classes; 2) while assessing peers, learners gain deeper understanding of what is expected of them, and they also improve their self-assessment skills.

- 'Teacher assessment" relates to 'non-standardized local assessment carried out by teachers in the classroom" (Leung, 2005:871). School teachers are usually confronted with a considerable assessment burden. They are expected to assess numerous individuals with respect to a considerable number of competences varying widely as concerns their degree of formality – from ongoing implicit assessment to regular syllabus-related testing all the way to local examinations forming the basis for crucial career decisions. At the same time, many teachers are only minimally trained for their assessment tasks.

  Many teacher assessments have a formative function as teaching and learning activities are adapted according to the assessment results. Teacher-based summative assessment remains important, even when standardized testing exists in a school system. While standardized testing often focuses on the results a system produces, teacher assessment is usually concerned with the

---

[8] Bachman & Palmer's book describes and illustrates in many other ways this core element of the approach adopted. It is not necessary to limit this, as they do, only to *language* assessment.

individual students in their class and school context. Teacher assessment therefore makes at least an important complementary contribution. The strength of teacher assessments lies in their flexibility and their adaptability to the contents actually treated in class; they can also consider particularities of the learning conditions and other relevant aspects. Moreover, teacher assessment is a suitable means to accommodate alternative forms of assessment, such as portfolio assessment or dynamic assessment building on multiple and frequent sources of information on the students.

Teacher assessment also has its drawbacks and problems. One of the main potential criticisms concerns the role and the far-reaching influence of the teacher in classroom-based summative assessment settings: teachers define the construct, produce and apply the instruments, interpret and apply criteria, decide about success or failure and even recommend the consequences to be derived. At the same time, quality requirements that apply to standardized testing do not normally apply to teacher assessment although some of the consequences concern basic career decisions and are therefore high stakes. Research has pointed to a number of aspects that could be improved:

- Teachers often are not sufficiently familiar with the constructs in question. The introduction of new concepts and materials relating to plurilingual and intercultural education should be used to familiarize teachers sufficiently through appropriate training and support. It is important to understand that construct-oriented assessment training may have positive effects on professional knowledge that reach far beyond assessment.

- Teachers often operate as individuals developing and applying their personal constructs. In order to effect improvement, they should be integrated in processes of group moderation with co-practitioners working on interpretations that are shared locally by everyone concerned. Integrated teaching of language and culture provides new opportunities to achieve such a shared basis. Training kits, optimally transparent instruments, intuitively accessible illustrations and other measures can help to build up local interpretations that are shared even more widely.

- High-stakes non-standardized assessments should be carefully monitored and amended or replaced when necessary. Some known effects such as the reference group error can be highly problematic when life opportunities are at stake: career decisions as they are regularly taken in school system at the entrance to higher levels of education should depend as little as possible on the incidental reference (sub-)group an individual belongs to. As a first step, the assessment instruments could be jointly produced, maintained and used by teams from different schools; as a further step, assessment experts could be called upon in order to help set up a professional testing system, and to make available instruments that are used uniformly on the whole reference population. In the latter case, care has to be taken that the impact of testing does not have undesirable consequences on teaching and learning.

- Expert assessment normally relates to assessment that is standardized to the highest degree that is appropriate and possible in a given situation. Experts are expected to be transparent about their claims to quality, to provide rationales and corresponding evidence. The Council of Europe's *Manual* on *Relating Language Examinations to the CEFR* (Council of Europe, 2009) and Bachman & Palmer's *Language Assessment in Practice* (2010) both provide guidance on best practice regarding the development and use of communicative language assessments. Many of the standards and measures suggested there can be adapted to other applications in the assessment of plurilingual and intercultural education.

Assessment experts are needed not only in the development, use and maintenance of *standardized* assessments but for any type of assessment. It should be clear from the above that self-assessment, peer and teacher assessment are legitimate and important forms of assessment that can bring many advantages to plurilingual and intercultural teaching and learning, and which cannot and should not be substituted. However, these types of assessment have to be better understood through research and reflective practice, and for these also, instruments and training modules should be developed. This is where experts are needed.

3 Assessing selected competences particularly relevant to plurilingual and intercultural education

Plurilingual and intercultural education is basically concerned with all language and (inter-)cultural learning but is particularly interested in plural, transversal and integrative approaches to language and culture.

As mentioned in the first chapter, one of the main intensions of *plurilingual education* is to brake down artificial borders between languages that are deeply imprinted in the heads of many language users, teachers and assessors, and to encourage the integral use of the plurilingual repertoire most people actually have. In sub-chapters 3.1 and 3.2 two ways of communicating in the presence of more than one language at a time – mediation and polyglot dialogue – are briefly presented, and discussed under assessment aspects.

The plurilingual, intercultural and strategic repertoires learners have at their disposal from their various experiences and specific training lend themselves to transfer when new communication tasks arise. In sub-chapter 3.3 the assessment of reading and reading strategies is focused on with special regard to *intercomprehension and intercomprehension strategies*.

The main preoccupation of *Intercultural education* has to do with the integration of 'otherness' in learners' thinking and actions. In sub-chapter 3.4 some purposes, options, methods and *caveats* regarding assessment in this field are discussed.

The discussion of the four areas will start with an overview of actual uses of the competences focused upon ("relevance") and will then generally follow the logical structure outlined in figure 2.1, which should guide assessment development as proposed by Bachman & Palmer.

3.1 Mediation

Relevance

Communicative language activities of mediation carried out by a third person make communication possible between persons who are unable to communicate successfully with each other directly. In principle, the communicative intentions of the mediating person do not go beyond facilitating (or mediating) understanding between the others who wish to communicate. Exact interpretation and translation are very specific examples of mediation. Mediation language activities can be limited to either the written or spoken channel, or comprise both channels. Also, mediation may be necessary within one language or involve two or more languages. The knowledge of the languages involved in the mediated communication may vary a great deal from participant to participant. Very often, intercultural issues or the specificity of tasks and topics may add to more basic language problems.

The following are some concrete examples of language activities involving mediation (excluding full translation):

- The German host sister of a newly arrived Colombian exchange student helps convey basic information on preferences regarding food, daily routines, special interests etc. between the host parents and the exchange student, thereby using her knowledge of German, Italian and English.

- A student travelling in another language region helps a tourist to understand an inhabitant of that region who is trying to describe to the tourist the way to a pharmacy.

- A family member, who has listened to a voicemail recording spoken in a foreign language, summarizes the main points in an SMS message and sends it to his or her mother.

- Students of a CLIL class read different factual texts in foreign languages, taking notes in their main language of schooling on behalf of the colleagues working on the same project.

It is likely that in schools that favour plurilingual and intercultural scenarios supporting exchange activities and CLIL, and do not implement distinct boundaries between different language classes, or language and 'content classes', there will be much more room and need for mediation language activities. The same will probably be true for other settings as soon as code switching and mixing and the inclusion of several languages in one communicative event have gained greater acceptance.

Intended consequences and decisions based on assessments

Specific courses of study for professional translators and interpreters show that the skills involved in these activities of mediation are complex and demanding. As translation professionals are expected to be versatile and precise, the requirements are generally very high.

However, many uses of mediation skills in plurilingual settings can be imagined that are less demanding. As is the case for other areas of communicative language use, intermediate levels of competence or competences relating to specific areas of use could be assessed and certificated, providing such areas and levels are sufficiently defined and described (preferably within a framework). As soon as mediation skills become part of language education syllabi, corresponding assessments for formative and summative purposes become an issue.

As mediation skills are useful and relevant in many real-world contexts, specific papers or modules relating to mediation skills could be added to existing language diploma examinations. This will be easier to do if the activities of mediation that are integrated remain in one language. Another option which might prove quite acceptable could be mediation between the main language used locally and the foreign language otherwise assessed in the examination. After all, the choice of languages to be included in an examination should be determined by the assessment and reporting needs that actually arise in a given context. In some specific-purpose contexts, the inclusion of mediation tasks in existing assessments seems particularly useful. Students from other language backgrounds entering tertiary education, for example, often have to give proof of sufficient competence in the local standard language. If needs analyses were conducted, it would probably become clear in many cases that partial skills in other languages, often English, as well as the ability to carry out activities that involve different languages, e. g. group work in mixed-language groups, or writing a literature review on the basis of a multilingual text corpus, are of crucial importance in higher education institutions. Therefore, an entrance examination oriented towards the actual communicative language needs of students at an institution might not only assess skills in more than one language but also mediation skills.

Construct to be assessed and interpretations

In terms of the CEFR, mediation is one of the four fundamental modes of language use beside reception, production and interaction. Mediation not only involves reception and production but also a specific 'mediation skill' which makes the link between the reception side and the specific conditions on the production side, i. e. mainly the addressee's characteristics. Mediators need to be able to establish what kind of language and communication can be understood by the addressee, and they should have the means to adapt (accommodate) their communicative behaviour accordingly. Depending on the interpersonal constellation, intercultural competences will have a more or less important share in this.

Language use in mediation spans a range of quite fundamentally different types of language activities: first of all, mediation within one language or variety can be distinguished from mediation involving more than one language. Secondly, groups of language activity can be identified through the channels (spoken/written) involved, the interactive or non-interactive character, as well as the order in which the channels are used in cases where both channels are involved.

In the *HarmoS* project in Switzerland (development of educational standards), a first attempt was made to include mediation[9]. For this purpose, a series of mediation language activities were described and attributed to seven categories of mediation language activities defined by three criteria. The criteria are as follows:

1   The languages involved: the language of the text heard or read (= source text) and the language in which the information is passed on (= target text)

   a)   Source text in the foreign language – target text in the L1 or local language of schooling

   b)   Source text in the L1 or local language of schooling – target text in the foreign language

   c)   Source text in the foreign language – target text in another foreign language

---

[9] The mastermind behind this proposed classification was Günther Schneider, University of Fribourg, Switzerland.

2   Spoken – Written: the source text as well as the target text can be either spoken or written. This yields the following constellations:

   a)   Spoken – spoken (e. g. informal interpretation between two individuals; pass on in spoken form the content of the latest news broadcast)

   b)   Spoken – written (e. g. leave a note in another language in order to pass on a message received on the phone)

   c)   Written – spoken (e. g. say in another language what is written on a sign)

   d)   Written – written (e. g. note down important information for somebody else from a website that is written in another language)

3   The communicative setting: the difficulty of the mediation activity varies according to the following factors: presence or absence of the persons whose utterances are to be mediated; immediate mediation or delayed mediation. The following constellations are relevant:

   a)   Face-to-face situation in which the mediating person as well as the persons who cannot understand each other directly, are all present. In this classical constellation for informal interpretation, asking for clarification regarding the source text, asking for support in formulating the target text, etc. are possible.

   b)   Face-to-face situation in which the mediating person passes on to an interlocutor, in the language of the interlocutor, foreign-language information both have heard or read (e. g. an announcement); asking for clarification is possible

   c)   Interpretation for groups, in front of a class or an audience; in these cases, asking for clarification is less possible.

   d)   Without direct contact; communication delayed in time (situational embedding, asking for clarification, or asking for help with formulations, etc. is not possible)

On the basis of the classification, the following seven types of mediation activities were suggested in the *HarmoS* project:

a)   mediation: oral ↔ oral: alternating between one foreign language and the main language of schooling (informal interpretation);

b)   mediation: oral ↔ oral: alternating between two foreign languages (informal interpretation);

c)   mediation: oral → oral: from the main language of schooling to a foreign language;

d)   mediation: oral → written: from the main language of schooling to a foreign language;

e)   mediation: oral → oral/written: from a foreign language to the main language of schooling;

f)   mediation: written → oral/written: from a foreign language to the main language of schooling;

g)   mediation: written → written: from a foreign language to a different foreign language.

An attempt was also made to formulate level-specific descriptors of communicative language proficiency in mediation, emulating the illustrative descriptors in the CEFR (mainly chapter 4). Due to the lack of empirically validated descriptors so far, elements from existing descriptors of receptive and productive language activities were exploited. This transfer from existing descriptors is of a hypothetical character.

The following example is a descriptor pertaining to category a), describing competences at level A2.2:

> Can reproduce in a simple conversation between a person who only speaks the language
> of schooling and a person who only speaks the foreign language, the gist of short, simple
> questions and information relating to personal detail (e. g. origin, hobbies) or everyday
> situations, provided the person speaking the foreign language uses clearly articulated
> standard dialect; the limited range of linguistic means in the foreign language may cause
> problems in formulating.

Subsequent empirical studies should explore the impact of the concurrent use of two languages and the need to operate in sometimes complex constellations, on the ability to perform 'as usual' in the

languages involved. It is to be expected that the increased cognitive load does have an impact. In addition, it would be useful to explore what role in describing mediation competence the use of strategies or techniques should play and which mediators use to make themselves understood when the recipient of a message has problems understanding the language(s) used for mediating: use of gestures, repeating and rephrasing, clearer and simplified speech, checking for understanding etc.

Mediation obviously covers a broad range of language activities involving various combinations of interpersonal settings and languages. Although there are specific skills of being an intermediary involved in activities of mediation, it does not seem to make much sense trying to describe and assess the competence of mediation in a general manner or as a whole. A more promising path is to look at the skills of mediation depending on what languages are involved, to what degrees they need to be mastered receptively and productively, while also relating to the different mediation types such as the seven identified above.

Assessments and assessment records

From the point of view of assessment, mediation tasks and their execution tend to include more variables and are often harder to interpret in terms of language competence than other language activities, especially those involving one language and one channel only, such as listening, reading or spoken production in one language. Some examples of this increased complexity follow:

- Mediation in live interactions depends heavily on the interlocutor's competence and behaviour. The need to adapt (*accommodate*) one's language flexibly and on the spot will vary considerably subject to the complexities of the input and the listeners' competence. These factors are hard to control for, or even to determine. In high-stakes assessments, role-plays involving trained interlocutors may help standardize the interlocutors' behaviours to some degree. However, processes of interpersonal *accommodation* in communication develop at a mostly unconscious level and may not be triggered realistically in simulations.

- The rating of performances poses additional challenges. For example, it remains to be defined (and made commonly known) what 'good' or 'not as good' *accommodation* to limited-proficiency beneficiaries of mediation means, what 'good' *foreigner talk* is, how the use of paralinguistic means and gestures is to be appraised, etc.

- Many mediation tasks that are not immediately interactive involve listening and/or reading of texts, processing, and then reproduction in an adapted form according to the needs of the recipient. For example, technical information from a manual needs to be explained in everyday language. These tasks raise questions that are quite similar to those encountered with live interaction. Full translation of a text in a corresponding style is quite straightforward: the task is easy to describe to an assessee, and rating criteria can be found that are quite intuitive. The process becomes more complicated when a task description must contain precise information on the modified text to be produced in order to suit the needs of a specific addressee of the text. The more formal an assessment, the less room there is for ambiguities in task instructions and task negotiation.

- A further potential disadvantage of such reception-processing-production tasks is that they are fairly complex and long and, at the same time, hardly transparent. From the results or product(s) of such a task, it may not be clear which parts were easily mastered and which caused problems or even led to a complete breakdown in communication while others worked well. This is true mainly for external assessments. In a pedagogic context, however, additional steps can be taken to make the learners reflect on and explain what happened 'inside' a task so that the assessment can be used as a crucial element in a formative cycle.

As indicated above, it makes little sense to attempt to cover a construct with so many aspects as mediation by means of assessments, or to make generalized conclusions regarding an underlying competence of mediation.

In many assessments in the field of language training for adults, conclusions relating to an underlying repertoire are of little importance in any event. Rather, the ability to accomplish well-defined tasks that are actually relevant at the workplace is emphasized. Testing for specific purposes often takes the so-called *work-sample approach*. Typically for this approach, real-world mediation tasks that are crucial for a specific work environment are selected and transformed into assessment tasks. Corresponding rating

criteria are often determined through the collaboration of language assessment specialists and (lay) stakeholders from the workplace. Summative work-sample testing is mainly interested in the degree of overall communicative success and much less in detailed diagnoses relating to the various task elements or even the language resources.

In educational contexts other than translator and interpreter training institutions, mediation skills are seldom assessed (the tradition of using translations as a means of learning and assessing grammar and vocabulary is not a counter-argument because of its different focus). Moreover, in the area of mediation much conceptual and empirical work remains to be done. This should not, however, impinge on the many opportunities that exist for mediation language use in plurilingual and intercultural education – indeed the contrary should be the case. Explorative and reflective teaching and assessment of mediation skills in innovative classrooms could help in developing descriptions of mediation language use in relevant constellations at different levels. Mediation and its assessment need to be practiced before scaled descriptors of mediation language activities, similar to the language proficiency descriptors in the CEFR, can expect to be developed.

3.2   Polyglot dialogue

Relevance

In institutional contexts, oral exchanges among speakers of different linguistic and cultural origins can be governed by different language regimes. The most common – and the most frightening, at least for some advocates of linguistic diversity – is the *lingua franca* regime, which is based on one common language only which is used by all participants. This language may or may not be the dominant or 'native' language of any of the participants. If it is, we speak of the monarchic model. However, other regimes beside the *lingua franca* model are also common in actual practice.

Polyglot dialogue is an interactional regime that allows for the use of two or more different languages or distant varieties in interpersonal interaction. Most often participants use one of their best-mastered languages productively and are capable of understanding the languages used by their interlocutors. The languages involved may be genetically closely related (as for example in the Germanic parts of Scandinavia or on the Dutch-German border on the Lower Rhine), or they may be less closely related but traditional languages of adjacent or even mixed territories (as on the German-French border in Switzerland). Polyglot dialogue thus involves the use of two or more different languages/varieties in oral interpersonal exchange in production – and consequently or reversely the use of two or more different languages/varieties in reception as well. The prerequisites for it to be efficient are that

- participants accept (explicitly or implicitly) that dialogue can take place without agreeing on one single language;
- participants are sufficiently competent to produce language in their strongest/stronger language in a way that is intelligible for the listeners present who master this language to a lesser degree;
- participants have sufficient receptive knowledge of the languages involved in order to understand their interlocutors using that language productively;
- participants have enough intercultural competence to manage potential or actual misunderstandings that may spring from the different language and cultural backgrounds. (This last point also applies to communication by means of a *lingua franca* to some degree).

Polyglot dialogue does not preclude code-switches by the speakers (insertions or longer stretches of the listener's preferred language) in order to assure comprehension of important utterances. It can involve language mediation as outlined in 6.1. Polyglot dialogue can be the rule in specific institutional settings or it can be part of ongoing negotiation of language regimes. It may be practiced in formal as well as informal settings.

The polyglot dialogue is an instance of plurilingual practice that has several advantages compared to monolingual regimes:

- It valorises languages/varieties that would never manage to become a *lingua franca* in a given setting.
- It allows speakers to express themselves in a language they feel most comfortable with, without excluding the use of other, less well mastered languages within the given setting.

- It presents an opportunity to develop the plurilingual repertoires of those present having, at the same time, little negative impact on the relevance of a dialogue.

For these reasons, the polyglot dialogue is an instance of plurilingual competence and proficiency 'at work' that seems to bear promising potential for maintaining and enhancing linguistic diversity within institutional contexts. Moreover, it illustrates the communicative potential of quite unbalanced language skills, i. e. the value that resides in partial mastery of languages (e.g. good receptive but much weaker productive skills).

Experience in Switzerland, for example, over many years has shown that people often are reluctant to engage in polyglot dialogue although it has been promoted by educational policy makers and in concept papers as a suitable and obvious way to communicate in a country having several official languages. The polyglot dialogue needs to find its way into the plurilingual classroom in order to get learners acquainted with this option and to demonstrate official support.

Intended consequences and decisions based on assessments

There are many possible arguments that can be advanced in favour of assessing and documenting competences that underlie and allow for successful interaction in polyglot dialogue - ongoing formative and summative assessment in school for pedagogic decisions at various levels being only one of these. Inclusion of polyglot dialogue in assessment rituals may add to its being recognized as a worthwhile option for communication in plurilingual settings – which, to date, is hardly the case.

Multinational companies that either explicitly or implicitly apply multilingual language regimes as well as services dealing with clients from various language background (in tourism, administration, etc.) have an interest in hiring employees – also at lower hierarchical levels – with these kinds of skills. For example, it might be important for an Anglophone pharmacologist, working in a branch of a European pharmaceutical company located in a Francophone territory, to be able to follow informal or even formal meetings that are held (partly) in French, to understand French-speaking interlocutors on the phone, etc. If polyglot dialogue is an accepted option, the pharmacologist can carry out these tasks even when his/her French productive skills are modest to nonexistent in the beginning.

The documentation and certification of language skills that are functional in and for polyglot dialogue adds value to an individual's portfolio because they represent resources that are sought after on the globalized or local job market (at least in multilingual countries).

From a language-learning perspective, it must be added that the receptive skills that are used and further developed in polyglot dialogue are also an important point of departure for the development of productive skills in the respective languages (see section 6.3 on intercomprehension for a more detailed rationale). Therefore, it seems reasonable, both from the perspective of language learning economy and in order to enhance the opportunities on the job market, to develop and document these competences in plurilingual educational settings.

The assessment of skills involved in polyglot dialogue makes sense for decisions taken by
  a) the plurilingual individual her/himself, e. g. with respect to new opportunities at the workplace, or regarding decisions about future language-learning needs or taking on new challenges at the workplace in divisions where different languages are used;
  b) teachers, particularly those preparing and counselling pupils who are to find a suitable place for vocational training;
  c) human resources managers in multilingual companies or services who need to determine the language skills of applicants and employees.

The documentation of skills in polyglot dialogue can be a particularly relevant addition to a person's portfolio. One of the tangible goals of this kind of documentation can be access to more and better job opportunities in multilingual companies and services.

Construct to be assessed and interpretations

In assessments of competences in polyglot dialogue, the emphasis is on the receptive skills, above all listening comprehension of formal and informal speech. The ability to follow oral exchanges in one or several other language(s) needs to be focused on. The listening-in-dialogue construct can consist of different types of comprehension of spoken language, ranging from global comprehension of what the

discussion is about, to detailed comprehension of contents and speakers' stances, including humour, irony, etc.

A further important competence to consider is active participation in dialogue whilst having the 'privilege' of using a language productively that is usually mastered at higher levels of proficiency than the language(s) that should be understood. Using a strong language for production and interaction - while for others this language poses problems - requires *accommodation* skills, i. e. the ability to gauge one's contributions precisely to the understanding the others demonstrate. Slightly different aspects of competence seem to be involved when it comes to joining in plurilingual dialogue at a micro level, e. g. closely interlocking one's turns with those taken by others in different languages; having a share in communication rituals that may not function the same way in the different languages used, etc.

Polyglot dialogue is often interspersed with *code switches* (symbolic, topic-induced, to meet productive and anticipate receptive needs) that seem to follow quite sophisticated rules that, however, are not codified. These switches are often used more or less consciously with social objectives in mind, such as identifying oneself as a member of certain groups, demonstrating solidarity, enhancing group spirit, etc. For consideration as a feature to be assessed, it is inevitable that 'good' code-switching is operationalised at least to some degree – less so for informal assessment, much more for standardized and high-stakes assessment.

There are instances of polyglot dialogue in some contexts where none of the languages used productively will be understood well enough by those who listen. In such cases, individuals may flexibly use chunks from other languages deemed useful (those spoken by others involved in the dialogue or a *lingua franca*) in order to compensate for the deficits on the reception side that might end a dialogue if they could not be overcome. This situation can easily occur, for example, when young people from different language regions meet, each of them still having a rather limited language repertoire. If they want to communicate with each other – in order to get to know each other, to agree on things to do, to find a solution to a practical problem, etc., it is likely that they will deviate from the prototypical polyglot dialogue and take refuge in any kind of language and strategic resources they have.

Assessments and assessment records

Assessments of polyglot dialogue should most usefully relate to two overlapping constructs: listening to turns and interactions in one language and more than one language simultaneously; interacting in such a dialogue.

Important aspects of listening comprehension in polyglot dialogue can be assessed using techniques from conventional communicative language testing: on the basis of audio or video recordings of dialogue sequences, test items have to be answered (cf. the 'understanding a native speaker interlocutor' scale in the CEFR). In a reduced form, an input sequence would contain one language in one turn only. In a more elaborate form, dialogue sequences featuring interaction between speakers of different languages would be used. As is usually the case in testing, the more complex the tasks are the more difficult it is to diagnose exactly the contribution of different factors, in this case different languages, for example. However, assembling complete tests that achieve sufficient coverage of a construct such as "understanding polyglot dialogue in three languages about everyday matters" seems possible. It would be interesting to study whether test-takers perform differentially in such plurilingual tests compared to tests that assess the languages involved consecutively.

Assessments of the ability to participate in polyglot dialogue that focus on listening comprehension only and use methods that do not allow the assessee to participate neglect one aspect: of the negotiating of meaning. Participants in non-hierarchic dialogues usually have the chance to ask for repetition or clarification when they need it. Therefore, listening comprehension tests of polyglot dialogue alone are not authentic as they neglect a crucial feature.

A higher degree of ecological validity (i. e. more authenticity) could be achieved through simulations of actual polyglot dialogue settings. A group of test-takers each having a different 'strongest' language would be confronted with common communicative tasks to be carried out. All participants would use their strongest language in the process. Modifications of this setting could include instructions to use code-switching in order to achieve certain goals or even to use other languages as far as the situation and the other participants' competences require.

The challenges involved in such settings when it comes to standardization, are quite obvious and imponderabilities abound. Problems might start with the need to have participants that are competent speakers of the languages that it is intended to include. It would then be quite difficult to brief the participants in a way that would assure some control over the course of the dialogue; they would probably be able to act even less according to their brief in interaction. This is a problem mainly because each participant also functions as an indispensable input source for the others.

A costly solution might be to have groups of trained interlocutors following a scripted plot; one candidate at a time would be inserted in this group and confronted with challenges in a strictly controlled manner.

If code-switching and accommodation in such a dialogue were to be assessed, additional challenges would arise because authoritative frameworks of these aspects do not exist.

To summarise, standardized testing on the basis of such ecologically valid tasks in polyglot interpersonal settings are quite a challenge while assessments focusing on listening comprehension seem to be feasible.

While we can see many potential difficulties that standardized assessment of competences in polyglot dialogue might encounter, we are much less sceptical as concerns more informal assessments of the same competences in educational contexts. For formative and low-stakes summative purposes many insights can be gained using reduced scenarios and functioning on hypotheses. Instead of having strong speakers of several languages, learners are given corresponding roles: one or two may use the main language of schooling, two others both the foreign languages learned, and perhaps a family language of a pupil can also be included. The objective of such an activity would simply be to show that a learner is able to operate in such circumstances, to activate a varied and already rich repertoire of knowledge, skills and strategies.

If code-switching as such is to be assessed in school, pupils could, for example, either watch code-switches on film or perform them themselves, and then give their interpretations or explanations, respectively. In this way they could show to what extent they understand possible rationales underlying code-switches

The portfolio concept is also relevant. Many people now have frequent plurilingual encounters. If they know that in order to get things done it is completely acceptable and brings many advantages to rely on one's strongest language as a first step, and then to mobilize the whole repertoire of language-related competences if necessary, learners may engage in such encounters more frequently. As is the case with intercultural encounters more generally (see 6.4), reflecting on and documenting such encounters on the basis of some guidance (different evaluation tasks and rubrics, self-assessment descriptors, etc.) provided in school or in a structured portfolio could have a formative effect and, at the same time, help to compile valuable evidence of competences and experiences relating to communicative settings such as polyglot dialogue and plurilingual dialogues more generally.

3.3   Intercomprehension in reading

Relevance

Reading competence is one of the most important sub-skills in education today. The PISA project (OECD) systematically monitors reading skills in the languages of schooling across a great number of countries. For many young Europeans in heterogeneous contexts, reading in the language of schooling means reading in a second language.

Probably all individual plurilingual profiles are asymmetric in nature, i. e. very often oral skills or receptive skills are better developed in an individual than the others. As far as literacy is concerned, receptive skills are much easier to develop than productive skills. Given the high investment necessary for the production of literacy in foreign and second languages, a focus on the receptive competences in literacy seems to be an important step towards a more up-to-date conception of the functional multilingual repertoire. Being able to read authentic texts in a variety of languages is a valuable skill in many fields, such as the Arts, Humanities or Social Sciences - for example in order to be able to read classics in their original language instead of translations. In professional life, receptive competences in foreign languages can be an asset that enables individuals to carry out all kinds of tasks that involve the collection, analysis and transformation of information from various sources.

Due to the usefulness of having reading skills in a variety of languages, methods have been developed in several *intercomprehension* projects[10] that are intended to facilitate learners' access to written texts in languages they have not expressly learned but which are genetically and thus often typologically related to languages already present in their plurilingual repertoires (mother tongues, foreign or second languages learned and taught in school, etc.).

Similar to the ability to engage in polyglot dialogue, intercomprehension skills facilitate access through reading to languages that might normally not be learned. Such partial skills are an important basis for the valorisation of 'smaller' languages, which means, in a European context, most notably, languages other than English. This is considered a desirable effect from a language policy point of view.

Intended consequences and decisions based on assessments

The assessment of intercomprehension skills is mainly relevant in school and language training contexts. Ongoing formative and summative classroom assessment will provide learners and teachers with feedback on deficits and the progress made. Including intercomprehension activities in regular assessments gives them the prominence that might be necessary to motivate learners (and teachers) to take these new developments seriously.

The valorisation of reading skills as such in whatever language follows an honourable European tradition. The learning of classical languages (Greek, Latin) has traditionally focused on reading skills almost exclusively. Putting emphasis on this purpose might, in the long run, contribute to a new perspective on the inter- and plurilingual potential of classical languages.

Construct to be assessed and interpretations

Developing intercomprehension skills relies on the development of inferencing techniques. In reading (or listening) to a text in an unknown but phylogenetically related language, multilinguals draw on their repertoire, including all potentially relevant languages/varieties as well as the common underlying literacy skills. All these elements or cues are potentially useful resources and transfer bases from which hypotheses may be formed with respect to cognate words, derivational patterns, parallel grammatical structures, informational and textual patterns and, finally, the meaning of sentences and texts. Interlingual inferencing in intercomprehension is therefore an extreme case of hypothesis generation based on new stimuli. The skill to draw such inferences and to generate a hypothetical grammar of the only 'halfway' unfamiliar language is a valuable resource for the life-long learning of languages. The construct to be assessed involves different aspects:

1) ability to quickly discover regularities within the unknown language and to relate them hypothetically to regularities known from other languages (e. g. in the domain of shared patterns of derivational morphology in the Germanic languages, in the domain of clause structure across Romance languages, etc).;

2) ability to recognize internationalisms and, more generally, cognates across languages, i.e. apply perceptual tolerance in identifying words in an unknown target language;

3) ability to develop efficient heuristics for identifying crucial and less significant elements in an unknown text; the former need to be understood, but not the latter.

Points such as those listed in paragraphs 1 and 2 are systematically documented in the *EuroCom* frameworks. Meissner (2004) gives a comprehensive account of his model of plurilingual didactics that attributes high value to intercomprehension. Meissner's account is fully compatible with our view of plurilingual language learning and use. To sum up, the construct that forms the basis for potential assessments of intercomprehension in reading combines crucial strategies for reading (difficult) texts in any language using cognitive processes that are crucial for foreign/second language learning. Both are related to the autonomous use and learning of languages and are cognitively demanding. The assessment of intercomprehension in reading that taps into a construct such as that presented here is also a form of assessing general language aptitude (cf. also Meara, Milton and Lorenzo-Dus, 2001).

---

[10] See, for example, the method of the „seven sieves" proposed by the *EuroCom* projects (Hufeisen & Marx, 2007; McCann, Stein & Stegmann, 2002). There are other European projects, such as *Galatea (Dabène, 2003),* working in similar areas).

Assessments and assessment records

Most assessments focusing on intercomprehension skills as such will target knowledge at the meta-level. Depending on how much of the existing training programmes in this field learners have already gone through, they are confronted with more or less accessible texts in languages related to others they already know better - for example, Danish texts for speakers of Dutch and English, Polish texts for speakers of Croatian, etc. Applying intercomprehension strategies is mainly an analytic task. Assessors will either ask learners to do sorting, matching etc. tasks – e. g. with words featuring the same stem – and/or to explain at a meta-level what they are actually doing. Both types of activity are informative for assessment purposes and contribute to further learning in the field of intercomprehension strategies.

When actual reading competence in genetically related languages, developed through intercomprehension strategies, is an objective, general methods of reading assessment can be used. These are not addressed in this document, however.

3.4 Intercultural competence

Relevance

Intercultural encounters – face-to-face or mediated in one way or another – and the necessity to deal with cultural otherness are basic facts of life in most of Europe. Recently, their importance has been accentuated through increased personal mobility, globalization, culturally less homogeneous societies, as well as challenges to historically grown political and social values through newly arrived inhabitants rooted in other cultures. Attitudes, knowledge and skills relating to competent handling of intercultural encounters are often considered important. Schools in many European countries are beginning to place more emphasis to intercultural education, and employers acting in an international environment or across language borders expect intercultural competence from job seekers. Nevertheless, it must be admitted that developing intercultural competence in learners cannot be considered a generally accepted priority as yet. Many do not seem to be familiar with its concepts while some are plainly opposed to them.

Intended consequences and decisions based on assessments

The rationale underlying school-based assessment of intercultural competence is mainly pedagogic. The results of the assessments will often be used formatively. Teachers and learners will try to find out what has been learned and understood about cultures and intercultural phenomena, to what extent learners open up in intercultural encounters, whether they are able to learn from such situations, if they can act constructively or even mediate when intercultural problems arise, etc. Based on the results, modifications of teaching and learning will be planned as necessary in order to improve on shortcomings identified through assessment. To the extent that schools allocate more resources to intercultural education, economic as well as pedagogic considerations are likely to come into play, with assessments required to provide more than pedagogically useful evidence of efficiency and effectiveness.

In a work environment, employers will be interested in having reliable evidence through assessments of those aspects of intercultural competence in their job candidates or employees that are needed to meet specific challenges posed through collaboration in multicultural teams or contacts with customers or distant partners. Employment and training decisions will be made accordingly. Employees' interest in assessments (formative or certificative) of their own intercultural competence can be expected to develop as employers' requirements in the intercultural field increase.

Construct to be assessed and interpretations

Sections B 2 and C of the first chapter summarize the different kinds of competences that are taken to be the main features of intercultural competence in the Guide. Later in the first chapter, intercultural competence is characterized as a conglomerate of five *savoirs*. Together, they can be taken to define intercultural competence.

As Byram (2008:221f.) rightly points out, some of these *savoirs* raise more questions than others when it comes to assessment, most notably *savoir être* (attitudes like curiosity and openness in relation to cultures) and *savoir s'engager* (critical cultural awareness – of potentially *any* culture). The desire to

establish profiles or even measures of curiosity, openness or critical awareness poses serious challenges. One of these resides in the fact that the corresponding construct definitions are still in the making: dimensions and components need to be identified or defined, and paths of progression should be established more reliably. A crucial point in relation to assessing attitudes such as curiosity and critical awareness is that they represent critical European values that are not shared by many other cultures. There are even hints of contradiction inherent in these concepts as the following consideration illustrates: an individual from another culture who is not open or curious due to his or her cultural identity will inevitably do badly on an assessment of cultural openness or curiosity; conversely, candidates who are particularly open and curious towards cultures that condemn openness and curiosity will attain a high grade on a curiosity and openness scale. As a response to these inconsistencies, Byram suggests that so-defined intercultural competence should be labelled as a partisan but vital part of European citizenship, due to the need of these societies to pass on basic European values, instead of non-binding relativism, to the next generations.

The other *savoirs* – cultural knowledge, skills of interpreting and relating cultural phenomena, skills of (inter-)cultural discovery and interaction – seem less problematic as constructs underlying assessments. However, even here, fundamental issues need clarification. First of all, there is a need to give more thought to the question of who the social groups are whose cultures are focused on. Is it, for example, the French culture of the population of France? Of which groups in France, then, speaking which language(s)? Even the high-prestige "defining culture" of France would be a fuzzy concept. In any situation that involves the construct of cultural otherness, even the most well-meaning proponent of intercultural competence cannot avoid but homogenizing, erasing and essential sing tendencies that often lead to problematic classifications of groups and individuals. The issue becomes even more complex when instead of nation, language or other large-group-based cultures the idea of individuals having intercultural identities, each developed on the basis of their personal biographies, enters the picture. In any case, defining the contents of a cultural-knowledge syllabus is not a straightforward task, and there is an inherent risk of codifying stereotypes whose real-world basis is in the process of decomposition. Therefore, extreme care has to be taken when the construct of cultural knowledge is defined for assessments.

In a work context, it is quite likely that only selected aspects of intercultural competence are of interest at any particular point in time. Needs analyses in the domains of use may help to target the constructs to be assessed.

As mentioned in the first chapter, *detailed* and *empirically validated* reference frameworks and corresponding assessment instruments, such as assessment tasks, competence descriptors and scales, are not yet available for the approach to intercultural competence used in the Guide. Interpretations making the link between observations based on assessments and the constructs they are supposed to be related to are still tentative (cf. the INCA grid). They can be used on that basis but interpretations have to be made with care.

Assessments and assessment records

Byram (1997) rightly draws attention to the fact that some competences can be better assessed over time while for others, assessments that take place at one given point in time can provide useful information. Also, some competences can be assessed by teachers or experts while for other competences, self-assessment or cooperative assessment are preferable.

Assessment over time is possible under the condition that the assesses are regularly available; this is the case for the learners themselves and, in school contexts, often peers and teachers. It is very likely that this kind of assessment does not usually take place under standardized conditions but rather informally, e. g. as updates to a learning log or portfolio or as teacher feedback sessions. These updates may be carried out periodically or motivated through specific teacher assignments or relevant experiences like exchange activities. The idea behind assessment over time is that repeated observation of an individual's feelings, thinking and behaviour in varied circumstances would finally give "the full picture", like the pieces of a jigsaw puzzle, and thus make up for the lack of formality standardized assessment would potentially provide.

Byram and others (Byram, 2008:224) prefer, and therefore recommend, portfolio-based assessment over other forms. Portfolio assessment is normally based on learner self-reflection and assessment, complemented by elements of cooperative assessment and assessment by others. Favouring self-

assessment has the double advantage that, on the one hand, it facilitates privileged insights, for example, into reflections and attitudes that are not easy to tap into otherwise; on the other hand, it helps develop self-reflection and self-awareness, which are both considered as key competences in intercultural education.

The Autobiography of Intercultural Encounters (AIE), developed by the Council of Europe's Language Policy Division as a specific addition to the Language Biography of the European Language Portfolio, is dedicated to the development of self-reflection and the assessment of intercultural competence at the same time. It is available in two versions: a standard version and a specific version for younger learners (up to age 11) adapted from cognitive, linguistic and emotional points of view. The AIE itself is complemented by a paper (in a short and a long version) on "Context, concepts and theory" for facilitators and advanced learners. The approach chosen is very straightforward: a series of questions and prompts is used to guide the learner's reflections on a single incident, I. e. an intercultural encounter. The questions provided reflect the various aspects of intercultural competence contained in the Byram (1997) model and help analyzing the incident and considering what learning may have taken place. The AIE may be used as a self-directed, cooperative or group activity. Through limiting itself to questions that help reflection and analysis, the AIE does not provide easily accessible explicit guidance on how further development might take place. It would therefore be beneficial from a formative perspective, except perhaps for very advanced learners, if the use of the AIE were embedded in teaching or counselling activities that could help to actively make use of the reflections to identify the next learning steps.

Within the INCA project, a portfolio was produced in addition to other assessment tools for intercultural competence assessment at the workplace and a tentative framework with levels. Its structure is similar to that of European Language Portfolios (ELPs) for adults but adapted to intercultural learning. It also consists of the three parts Passport, Biography and Dossier. The *Passport* section is intended to be a summary of more or less formal assessments by others, as well as guided self-assessment based on INCA material. Parallel to the reference level descriptions in the ELP, the INCA *attainment grid* is used as a reference framework with dimensions and levels that creates coherence, transparency and comparability across the different elements of reference collected. The *Biography* section is intended as a personal record of intercultural experiences, bringing together details on relevant experiences, self-reflection on one's intercultural behaviour as well as "diary entries" of experiences during which intercultural learning happened. The *Dossier* part encourages learners to present personal evidence provided by themselves and testimonies by so-called witnesses, illustrating the assessment results recorded in the Passport and experiences recorded in the Biography.

The two portfolios just presented combine assessment at one point in time with assessment over time. The AIE limits itself to the analysis of single intercultural encounters. Through repeated analysis and reflection using basically the same procedure on different encounters, development of intercultural competence may be observed. The INCA project provides more versatile instruments that may be used to get a fuller picture.

Among the core elements provided by the INCA project are test tasks complemented by guidance for assessors. The task types are either scenarios (text or video-based) or role-plays through which applied competences relating to the different *savours* may be demonstrated. In the INCA tasks, the test takers are asked to comment either in spoken or written form on (critical) scenarios posing different types of challenges. The performances are then to be assessed by (trained) assessors based on the INCA *attainment grid.*

As repeatedly stressed, the AIE instrument and the INCA tasks and grid are tentative. That means, they are still subject to improvement and validation at different levels. In addition to testing institutions, schools and teachers are invited to contribute in their respective contexts.

Teachers/facilitators who are willing to experiment could transfer the INCA approach and instruments to their contexts. Being well informed on concepts of intercultural competence would certainly be a great advantage, as would also working in a team. The experiences made locally should be documented and reflected over time in order to take things forward and to develop better and more practicable procedures to assess intercultural competence, and to better understand actual progress made by specific categories of learners.

Testing experts would first be interested in determining through expertise and surveys what intercultural competence profiles in individuals are actually deemed important in (specific) real-world contexts, including education at different levels. They would then need to develop optimally valid assessments, accompanied by arguments and evidence for the degree of validity claimed. In addition to points mentioned above, such as the lack of sufficient empirical evidence to support proposed frameworks, one specific point is often highlighted as posing a major challenge to standardized testing of intercultural competence - the danger of *display behaviour* when it comes to attitudes. It is suspected that test-takers with sufficient training might just behave as if they were curious, open or tolerant in order to receive positive judgments in a high-stakes assessment context. There are at least two points worth considering that might mitigate this concern: 1) in psychological testing standardized instruments such as attitudinal questionnaires exist that are suitable to detect "dishonest" behaviour, so this problem could probably be minimized; 2) it is open to discussion whether assessments of intercultural competence need to – or are even authorized to – disclose the deepest motives of behaviour; instead, it could content itself with consistent evidence for the claim that the assessee is capable of performing in an interculturally competent manner in every new similar contexts (cf. also Lázár et al, 2007).

4 Conclusion

Educational contents normally get more visibility and recognition once they are assessed and included in decision-making processes. Assessment in a broad sense – including a whole range of assessment types for different purposes, from informal to very formal – also has the potential to motivate learners and to contribute to better, i. e. well-targeted, more profound, more efficient and more effective learning. Care must be taken, however, that the ancillary function of assessment in the teaching and learning process is not reversed. All this is also true for plurilingual and intercultural education.

The use(s) of any kind of assessment, no matter how informal, should be based on transparent reflection and argument; moreover, evidence should be provided that what is claimed to be assessed by an assessment, is actually being assessed. Otherwise, it is of reduced validity. Depending on the (social) consequences that follow from decisions made on the basis of an assessment, the validity requirements regarding the assessment are higher or lower respectively.

The explorations into the assessment of plurilingual and intercultural competences undertaken in this study, have shown that, at this point, many possible assessments are well-suited for formative purposes, especially when they involve the learners as self-assessors and allow for some imprecision and goodwill. With regard to standardized, high-stakes assessment, the possibilities are much more limited. In some cases, these limitations seem to be due to the nature of the object, in some others it seems that reliable high-stakes assessments could actually be devised as soon as reliable reference frameworks and testing instruments have been developed. What is needed most at this time, however, is actual practice: plurilingual and intercultural learning, teaching and assessment should to be put in place in innovative ways, and then be evaluated critically in order to make progress.

5 Bibliography

Bachman, L. (2005): Building and Supporting a Case for Test Use. Language Assessment Quarterly, 2, 1, pp.1-34.

Bachman, L. & Palmer, A. (2010). *Language Assessment in Practice*. Oxford; New York: OUP.

Beacco, J.-C. (2004). Une proposition de référentiel pour les compétences culturelles. In J.-C. Beacco, S. Bouquet & R. Porquier (eds.), Niveau B2 pour le français, pp. 251-287. Paris: Didier.

Beacco, J.-C. & Byram, M. (2007). *From linguistic diversity to plurilingual education: Guide for the development of language education policies in Europe,* Strasbourg: Language Policy Division, Council of Europe (www.coe.int/lang → Policy Instruments → Policy Guide) (18/08/10)

Byram, M. (1997). Teaching and Assessing Intercultural Communicative Competence. Clevedon: Multilingual Matters.

Byram, M. (2008). From Foreign Language Education to Education for Intercultural Citizenship.

CARAP – Candelier, M., Camilleri-Grima, A., Castellotti, V., de Pietro, J.-F., Lörincz, I., Meissner, F.-J., Schröder-Sura, A., Noguerol, A. (²2007). Framework of Reference for Pluralistic Approaches to Languages and Cultures. European Centre for Modern Languages (ECML) - Council of Europe. http://carap.ecml.at/ (29/06/10).

CDIP / Conférence suisse des directeurs cantonaux de l'instruction publique (2010). Standards de base pour les langues étrangères. Bern: CDIP. Internet: http://edudoc.ch/record/36465/files/Standards_L2_f.pdf?version=1 (30/06/10).

Coste, D., Cavalli, M., Crişan, A. & van de Ven, P.-H. (2009). *Plurilingual and intercultural education as a right*. Strasbourg: Council of Europe, Language Policy Division. (www.coe.int/lang: "The platform of resources and references for plurilingual and intercultural education": Box: 'The learner and the languages present in school') (18/08/10).

Council of Europe (2009). *Autobiography of Intercultural Encounters*. Strasbourg: Council of Europe, Language Policy Division. (www.coe.int/lang → Autobiography...).

Council of Europe (2009). *Regional, minority and migration languages*. Strasbourg: Council of Europe, Language Policy Division. (www.coe.int/lang: "The platform of resources and references for plurilingual and intercultural education": Box '*Regional…*') (19/08/10).

Council of Europe (ed.) (2001). *Common European Framework of Reference for Languages: learning, teaching assessment*. (CEFR) Cambridge: CUP. (www.coe.int/lang → Policy Instruments → Common European ...) (29/06/10).

Council of Europe (ed.) (2009). *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual.* Strasbourg: Language Policy Division. (www.coe.int/lang → Policy Instruments → *Common European* ... → *Relating language examinations*…) (29/06/10)

Dabène, L. (2003). De Galatea à Galanet. Un itinéraire de recherche. In C. Degache (ed.), Intercompréhension en langues romanes. Du développement des compétences de compréhension aux interactions plurilingues. De Galatea à Galanet. (=lidil 2003/28), pp. 23-29.

Dam, L. (1995). Learner Autonomy 3: From Theory to Classroom Practice. Dublin: Authentik.

Dickinson, L. (1992). Learner Autonomy 2: Learner Training for Language Learning. Dublin: Authentik.

Fleming, M. (2009). *The Use of Descriptors in Learning, Teaching and Assessment*. Strasbourg: Council of Europe. (www.coe.int/lang: "The platform of resources and references for plurilingual and intercultural education": Box 'Reflections on the use of descriptors in learning, teaching and assessment')

Holec, H., 1981: Autonomy and Foreign Language Learning. Oxford: Pergamon. First published 1979, Strasbourg: Council of Europe.

Hufeisen, B. & Marx, N. (ed.) (2007). EuroComGerm – Die sieben Siebe. Germanische Sprachen lesen lernen. Aachen: Shaker

INCA (Intercultural Competence Assessment) Project. A Leonardo da Vinci project. Online: www.incaproject.org (30/06/2010).

Janua Linguarum – Candelier, M., Andrade, A.-I., Bernaus, M., Kervran, M., Martins, F., Maurkowska, A., Noguerol, A., Oomen-Welke, I., Perregaux, Ch., Saudan, V., Zielinska, J. (2004). Janua Linguarum - The Gateway to Languages. The Introduction of Language Awareness into the Curriculum: Awakening to Languages. Strasbourg: Council of Europe. Internet: http://www.ecml.at/documents/pub121E2004Candelier.pdf (26/06/10).

Lázár, I., Huber-Kriegler, M., Lussier, D., Matei, G. S. & Peck, C. (2007). Développer et évaluer la compétence en communication interculturelle. Un guide à l'usage des enseignants de langues et des formateurs d'enseignants.

Leung, C. (2005). Classroom teacher assessment of second language development: construct as practice. In Handbook of research in second language learning and teaching (ed. E. Hinkel), pp. 869-888. Mahwah, N.J.: Lawrence Earlbaum Associates

Little, D. (1991). Learner Autonomy 1: Definitions, Issues, Problems. Dublin: Authentik.

Little, D., 2005. The Common European Framework and the European Language Portfolio: involving learners and their judgements in the assessment process. Language Testing 22 (3), pp. 321–336.

McCann, W.J., Klein, H.G., Stegmann, T.D. (2002). EuroComRom -The seven sieves: How to read all the Romance languages right away. Aachen: Ed. EuroCom.

McNamara, T., Roever, C. (2006). Language Testing: the Social Dimension. Malden, MA: Blackwell.

Meara, P. M., Milton, J. & Lorenzo-Dus, N. (2001). Language aptitude tests: lognostics. Newbury: Express.

Meissner, F.-J. (2004). Introduction à la didactique de l'eurocompréhension. In F.-J. Meissner, C. Meissner, H. G. Klein & T. D. Stegmann (eds.), EuroComRom - Les sept tamis: lire les langues romanes dès le départ. Avec une introduction à la didactique de l'eurocompréhension, pp. 7-140. Aachen: Shaker.

Messick, S. A. (1988). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed.), pp. 13-103. New York: American Council on Education/Macmillan.

North, B. (2000). The Development of a Common Framework Scale of Language Proficiency. New York etc.: Peter Lang.

O'Malley, J. M., Chamot, A.U. (1990). Learning strategies in second language acquisition. Cambridge: Cambridge University Press.

Oscarson, M. (1997). Self-Assessment of Foreign and Second Language Proficiency. In Encyclopedia of Language and Education, v. 7 (Language Testing and Assessment), pp. 175-187.

Oskarsson, M. (1984). Self-Assessment of Foreign Language Skills: a Survey of Research and Development Work. Strasbourg: Council of Europe.

Oxford, Rebecca (1990): Language Learning Strategies: What Every Teacher Should Know. New York: Newbury House.

Rosenthal, R. & Jacobson, L. (1992). Pygmalion in the classroom: teacher expectation and pupils' intellectual development. New York: Irvington Publishers.

Schneider, G., Lenz, P., Studer, Th.; Konsortium HarmoS Fremdsprachen (2009). Fremdsprachen. Wissenschaftlicher Kurzbericht und Kompetenzmodell. Bern: EDK. Internet: http://www.edudoc.ch/static/web/arbeiten/harmos/L2_wissB_25_1_10_d.pdf (01.03.10).

Shohamy, E. (2001). The power of tests: A critical perspective of the uses of language tests. London: Longman.