



UNIGE:
CRITERIA-BASED COMPARISON IN
LANGUAGE POLICY: □ PRINCIPLES AND
METHODOLOGY

MICHELE GAZZOLA
FRANÇOIS GRIN

ABSTRACT

Local, national and supranational contexts characterised by linguistic diversity usually require some form of public intervention to manage multilingual communication. Public interventions are typically carried out by the state or by its surrogates and they can affect a varying number of domains. Clearly, this does not exclude the possibility that other forms of management of linguistic diversity appear without the intervention of public authority, usually as a result of the decentralised action of agents (e.g. individuals, associations, etc.). Nevertheless, it is not possible for public authorities not to have any language policy, since at least one code must be used for ordinary administration and more generally in public education, media, etc. Hence, the problem of “choice” is at the heart of diversity management (which languages to use, to what extent, on which territory, etc.).

Our paper provides a framework for structuring language policy choices based on the theories and methods of policy analysis. The theory of policy evaluation and the theory of indicator design is revisited in the perspective of multilingual communication. This paper is divided in two parts. In the first section, we present how procedures of ex ante and ex post policy analysis can be adapted to language policy design, implementation and evaluation. The ex ante policy evaluation process is characterised by three distinct phases. The first phase can be named “problem analysis”, and it includes the understanding the problem to be solved (e.g. managing multilingual communication) and the choice and explanation of relevant objectives and constraints. The second phase, “evaluation design”, consists in defining the alternative plans to be compared, identifying and involving stakeholders, identifying evaluative questions and defining evaluation criteria, and choosing methods and tools for evaluation. In this paper we shall focus on two specific evaluation criteria, namely efficiency (interpreted as cost-effectiveness) and fairness, and on one particular evaluation method, that is, cost-effectiveness analysis. In the third phase evaluation is implemented. This phase consists of several steps, that is, analysing the internal structure of the policy plan (or plans) proposed,

assessing the policy implementation plan, examining or designing the information system, estimating outputs and outcomes, and providing final judgment and recommendations. Ex post evaluation shares most of the steps of ex ante evaluation, except for “problem analysis” which is clearly not relevant.

In the second section, we focus on a specific step of the third phase of evaluation, that is, the examination and design of language policy information system. This step consists in designing evaluative indicators or assessing indicators already used. The information system is populated with data and it is meant to provide valuable and reliable support for the policy maker. Indicators are built in relationship with the evaluative questions addressed. For this reason, we focus on the design of effectiveness and fairness indicators for language policy evaluation, and on the assessment of the desirable properties of these indicators, such as validity, sensitivity and reliability. Finally, we discuss the desirable properties of the indicator systems as a whole (coverage, balance, selectivity and relevance) in order to provide an overall assessment of the adequacy of the language policy information system.



1. INTRODUCTION: THE NEED FOR EVALUATION IN LANGUAGE POLICY AND PLANNING

Although the potential contribution of policy analysis to language policy and planning (LPP) was already clear in sociolinguistics in the 1970s (Jernudd 1971, Rubin 1971, Thorburn 1971), the literature on LPP has mostly been silent on evaluation until the 1990s. As Ricento has noted, “what has not been much discussed is the practice of language planning, that is, the development, implementation, and evaluation of specific language policies” (2006: 18).

However, the first contributions in the evaluation of language policy were published in the 1970s, as a result of the work of Canadian economists interested in analysing the advantages and the drawbacks of Canadian bilingual language policy (Vaillancourt 1983, Vaillancourt 1985). Nevertheless, in most cases the connection to language policy in these early contributions was somewhat derivative, and research often responded to specific developments of which authors had personal experience. The picture has started to change since the 1990s, when a growing number of researchers

coming from political science, rational choice theory and economics, started to import concepts and tools from policy analysis in LPP, showing that language policies can be viewed and analysed as a form of public policy.¹

The goal of this article is to propose a general methodology for the evaluation of language policies, focusing in particular on the concept of policy indicators. So far most of the contributions in language policy evaluation have focused on policies aimed at supporting regional or minority languages (cf. Jaime et al. 1995, Grin et al. 2002, Grin 2003, Simó et al. 2006). This article, by contrast, addresses the selection, design and evaluation of language policy in general (though we shall generally refer to it simply as “evaluation”), including policies aimed at managing multilingual communication in international institutions, multilingual states, etc. It is therefore intended to contribute to the development of a “culture of evaluation” in LPP, whose importance is increasingly emphasised among LPP practitioners and scholars.

Section 2 clarifies the role of policy analysis with respect to other research traditions in LPP. Section 3 provides a detailed presentation of the evaluation process, and it discusses some of the major problems related to adapting a standard policy analysis framework to LPP. Section 4 introduces and discusses the concept of policy indicator, focusing in particular on the assessment of the quality of indicators and indicator systems. Section 5 discusses different examples of existing linguistic indicators. Section 6 concludes.

2. LANGUAGE POLICY EVALUATION IN PERSPECTIVE

There is a logical sequence in the questions addressed in LPP, and some questions come first. The questions of “why” and “what”, that is, for what reasons a policy should be done and what should be done, come before the question of “how” to do it, as suggested in figure 1 which offers a bird’s-eye view.

The question of “why” is usually addressed in public debate. A particular language policy can be justified by several reasons, like political and ideological claims related to decolonisation processes, moral reasons (e.g. equality of opportunities), legally binding provisions, and so on. These reasons may be

¹ In this article, we use the terms “policy analysis” and “policy evaluation” interchangeably, although several differences exist between the two concepts (Geva-May and Pal 1999). Policy analysis is the object of a vast literature. Among others, cf. European Commission (2008), Rossi et al. (2004), World Bank (2004), Weimer and Vining (2005), and Knoepfel et al. (2007). On the application of policy analysis to LPP, see Grin (2003a).

intertwined and usually are. The discussion on “why?”, therefore, is the object of public debate which can be nourished thanks to the contributions of different disciplines.

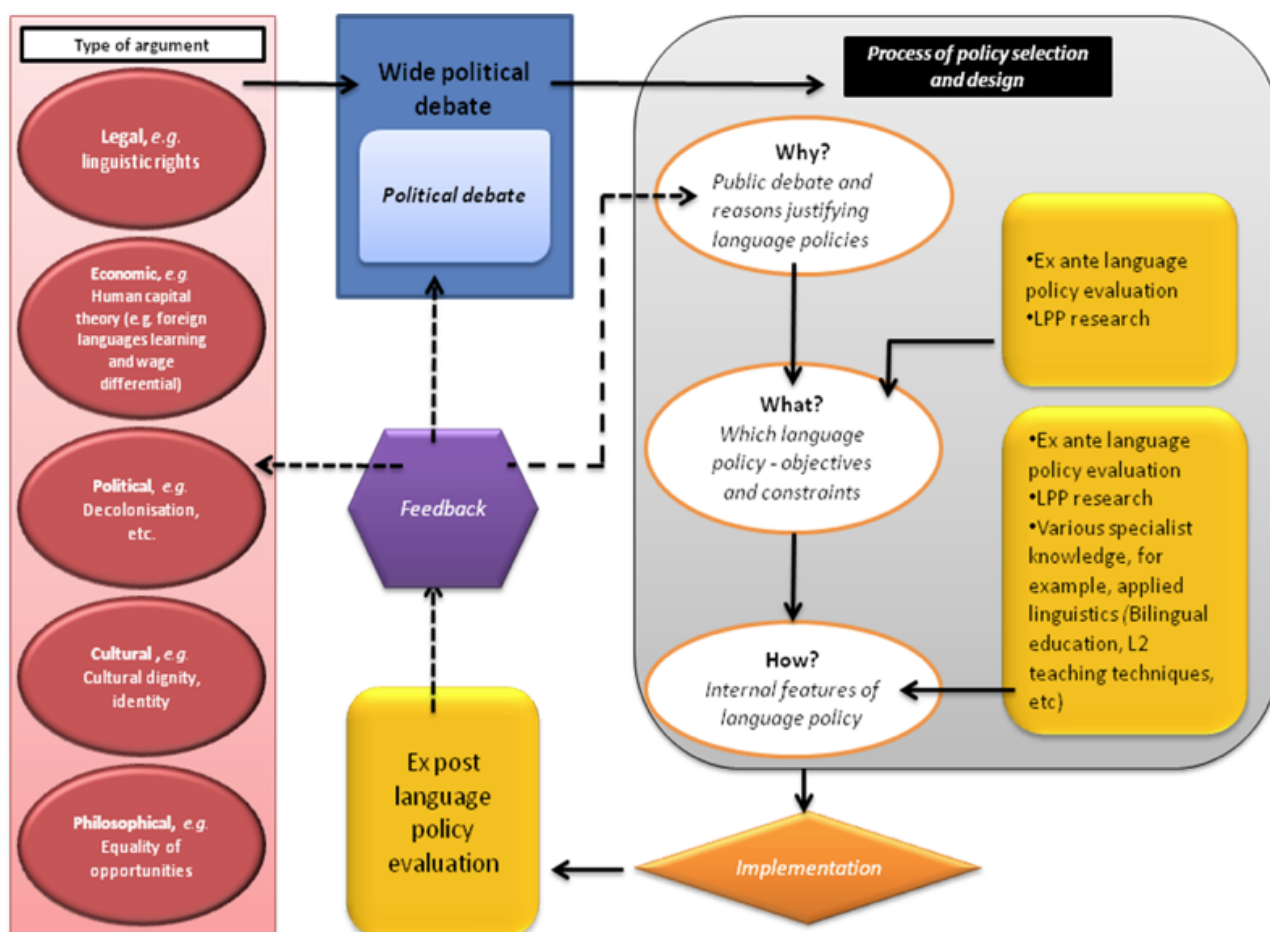


Figure 1 : The role of different approaches to language policy design

By contrast, specific approaches are requested to answer questions related to “what?” and “how?”. In order to answer the “what?” question, in fact, we have to put objectives in relation with constraints, compare alternative policy options, and finally evaluate them. Policy analysis and LPP provide precisely the methodology and tools appropriate for this purpose.

Policy evaluation is also useful for addressing “how” questions, but input from applied linguistics are indispensable for this purpose, since they help examine how specific language policies can be implemented in practice, which particular techniques are more appropriate for obtaining better learning performance, etc. For example, the implementation of a policy aimed at increasing foreign language skills can benefit from inputs provided by applied linguistics, such as

studies and experience concerning content-language integrated learning (CLIL). Likewise, support for a minority language usually requires some form of corpus planning, etc.

Policy analysis, therefore, serves two functions. It is used before a policy is adopted (ex-ante evaluation) and after it is implemented (ex-post evaluation). In what follows, we do not address questions related to the “why” of language policies, and we focus on “what” and “how” questions, and therefore on the specific features of the evaluation process.

3. LANGUAGE POLICY AS AN OBJECT OF EVALUATION

A public policy can be defined as, “a series of intentionally coherent decisions or activities taken or carried out by different public—and sometimes private actors—, whose resources, institutional links and interests vary, with a view to resolving in a targeted manner a problem that is politically defined as collective in nature” (Knoepfel et al. 2007: 24, emphasis added). Just like public policies are a response to “public problems”, language policies can be viewed as a response to “language problems”.² We do not address here the question of who are the actors involved in defining what a problem is, say, political parties, associations, media, etc. We take the broader view that in a given context public authorities are involved in addressing a language problem, such as increasing the average level of pupils’ foreign language skills, promoting a regional or minority language, choosing a language regime for an international institution, etc., and we focus on the procedure to guide their choices.

A distinction should first be made between ex-ante and ex-post evaluation. Ex-ante evaluation is prospective, and it is mainly aimed at supporting decision making and therefore choice. Ex-post evaluation is mainly retrospective, and it is used to assess the final results of the policy implemented, to clarify whether the results obtained are consistent with the initial objectives, and to provide a general assessment of why the policy has been managed. Policy results, in turn, are interpreted in the light of some evaluative criteria in order to deliver a well-argued final judgment on the policy.

This judgment provides feedback not only to policy makers, but also to stakeholders and taxpayers. Evaluation, thus, is also intended to increase

² According to Rubin, for example, language planning “focuses upon the solution to language problems through decisions about alternative goals, means, and outcomes to solve these problems” (1971: 218).

accountability. Policy evaluation, therefore, is not only aimed at passing a well-founded judgment on the effects of a policy and on its processes. It can also be interpreted as a process of organisational learning that helps to improve policies over time. Figure 2 shows a general framework for a step-by-step plan for the evaluation of language policies. Let us present these steps in more detail.

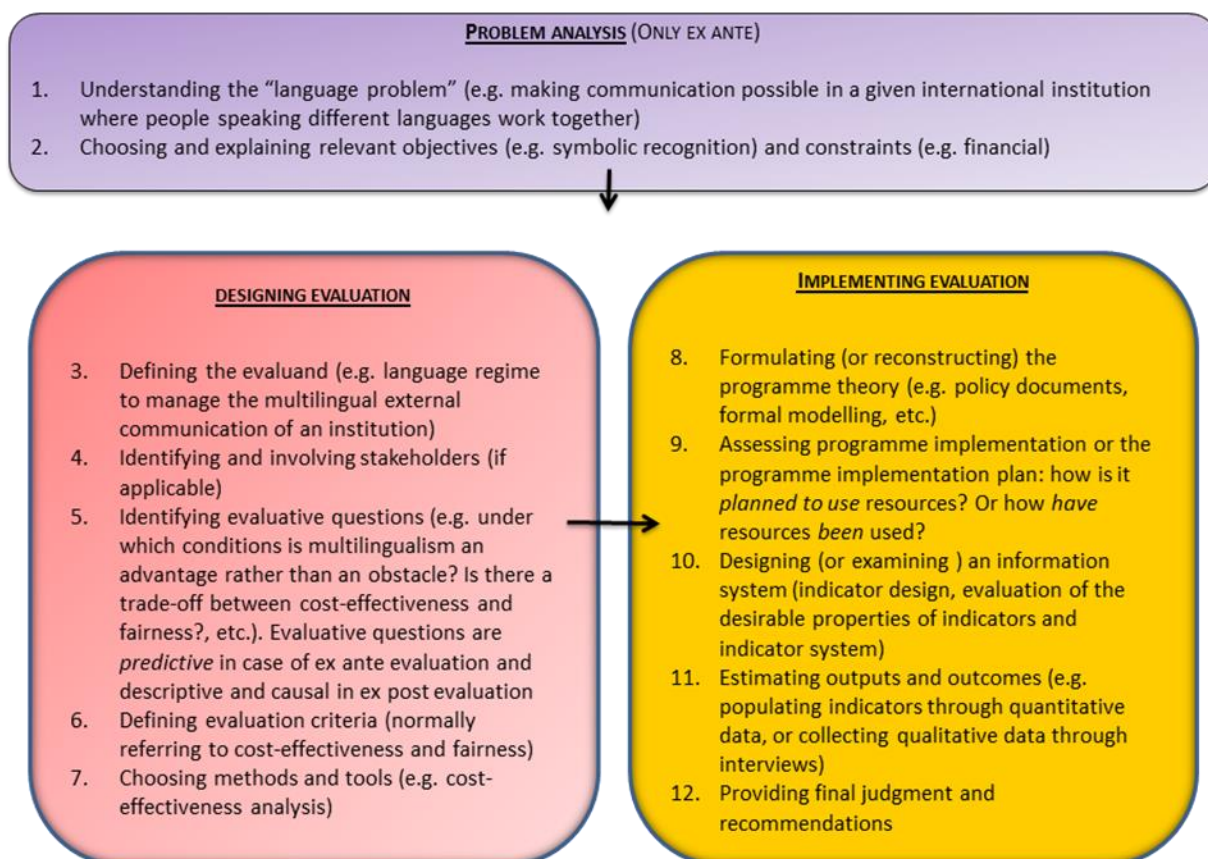


Figure 2 : Steps in language policy evaluation

3.1 PROBLEM ANALYSIS

The evaluation of a policy is usually preceded by a preliminary phase called problem analysis.³ During this phase, the need for a public intervention is assessed on the basis of a description of the nature and extent of the public problem to tackle, taking the political, institutional, economic and social contexts into account (Weimer and Vining 2005: 328). As noted by Palumbo (2001: 330), the political context influences policy priorities and objectives, the institutional context defines the conditions under which objectives can be pursued. In addition, the economic context determines the quantity and the type of resources available for implementing programmes, whereas the social

³ If not specified otherwise, no distinction is made between policy, programme, intervention and project.

context influences the perception and the definition of needs. The result of problem analysis is a set of general objectives and a definition of the target population.

In practice, problem analysis consists in clarifying which “language problem” is to be solved (step 1), for instance, making communication possible in a given international institution where people speaking different languages work together, choosing a set of measures to increase the vitality of a given minority language, improving the average level of proficiency in a given foreign language of pupils studying in a certain region, etc.

Understanding the problem requires a general assessment of the symptoms of the problem, a proper framing of the problem itself, and a clear identification of relevant policy variables. For example, if the problem at hand concerns the management of multilingual communication in an international institution, a correct characterisation requires a general picture of the linguistic competence of agents, the market prices for translation and interpreting, etc.

Step 2 consists in choosing and explaining the relevant objectives of a policy—e.g., increasing skills in a given foreign language, symbolic recognition of language communities, etc.—and constraints (e.g. financial). Problem analysis is followed by solution analysis, which can be divided in two macro-phases, namely, design and implementation.

3.2 EVALUATION DESIGN

The design of an evaluation is at least as important as the implementation of evaluation itself, since bad design will negatively affect implementation. Evaluation design can be structured in five steps, that is, (3) defining the evaluand, (4) involving stakeholders, (5) identifying evaluation questions, (6) defining evaluation criteria, and (7) choosing evaluation methods and tools.

3. Defining the object of the evaluation itself (or “evaluand”). The policy to be evaluated is usually delimited according to some relevant dimensions, for example, territory, type of institutions involved, period of time considered, etc. In ex-post evaluation, the evaluand is the policy already implemented. In ex-ante evaluation, the evaluand is the policy (or the set of policies) proposed.

4. Identifying and involving stakeholders. Identifying and involving relevant stakeholders has several practical advantages. First, it helps to obtain good quality information. Secondly, “evaluators along with programme managers have an interest in ensuring that there is ownership of evaluation findings” (European Commission 2008: 39). Since evaluation is not an end in itself, it has to be useful first to actors involved. The degree of involvement of stakeholders depends on the type of evaluation required. If evaluation is commissioned by an institution, the active involvement of stakeholders is usually necessary and advisable. Nevertheless, in other cases, for example when evaluation is carried out independently for academic research purposes, this phase may not be particularly relevant.

5. Identifying evaluation questions. A central component of evaluation design consists in identifying relevant questions, since often “the value of evaluation for policy makers lies as much in posing the right questions as in providing precise answers” (European Commission 2008: 36). Formulating evaluation questions also requires a deep understanding of the evaluand also from the point of view of decision makers and stakeholders (if involved). Formulating evaluative questions making sense for stakeholders implies exploring the semantic space of the evaluand and negotiating meanings with the persons involved.

Evaluative questions are of a different nature depending on which evaluation phase they refer to. Examples are “under which conditions is multilingualism in universities in a given country an advantage rather than an obstacle?”, “has language policy 1 proved more effective than policy 2 in increasing the vitality of a given language?”, etc.

In an ex-ante evaluation, questions will typically be predictive. In an ex-post evaluation, evaluative questions will typically be descriptive (e.g. “what happened?”) and causal (e.g. “to what extent is what happened due to the policy adopted?”). A proper process of formulation of evaluative questions clarifies what should be evaluated, and therefore suggests a set of instruments (indicators) in terms of which answers can be found (see section 4 below).

6. Defining evaluation criteria. Evaluation questions are related, at least implicitly, to evaluation criteria. Generally speaking, the main evaluation criteria in policy evaluation are relevance, effectiveness, efficiency and/or cost-effectiveness. Relevance is the appropriateness of the objectives of policy (or

policy options) in relation to the problems it is supposed to tackle. For example, if the objective is to improve skills in a given foreign language (say, X) for adults in a given region within a short period of time (say, 1 year), a policy aimed at increasing the exposure to X in secondary school is not relevant.

Effectiveness consists in the consistency between objectives and the outcomes (or results) achieved. At this point, it is important to recall the distinction between inputs, outputs and outcomes (cf. table 1 below).

Table 1: Policy inputs, outputs and outcomes

Concept	Definition
Input	Financial, human, material, organisational and regulatory means mobilised for the implementation of a policy.
Output	What is funded and achieved (or realised) through the resources allocated to the policy. E.g. number of participants in language courses (say, language X) subsidised by public authorities.
Outcome	Effects of the policy on the target population following the policy. E.g. number of participants who got the language certificate (say, B1-level certificate of knowledge in language X).

Source: adapted from European Commission (2008)

Outcomes can be of positive (“advantage”) or negative (“disadvantage”) for directly concerned actors.

For example, assume that the objective of a policy is to improve skills in a given foreign language (say, X) for adults in a given region within a short period of time (say, 1 year), through publicly subsidised language courses. The direct outputs of this policy are the number of participants to language course. Outcomes are the final effects of the policy on the public problem in question, that is, increasing language skills, or perhaps, even further downstream, the increase in the use and vitality of the language made possible by higher level of skills in the population. The policy, therefore, is effective (that is, it has an effect) if participant pass the final test certifying a level of skills in language X

equal to, say, level B1 of the Common European Framework of Reference for Languages.⁴

Efficiency and cost-effectiveness are two different concepts, but both refer to resource allocation. Cost-effectiveness denotes a technical relation between the inputs and the outcomes. A policy can be viewed as cost-effective if a certain outcome is achieved at the lowest possible cost in resources (or equivalently, given a certain cost, the results are as good as possible). Cost-effectiveness, therefore, refers to a comparison between the costs of a policy and its outcomes, where the latter are not expressed in monetary form, but in some other unit of measurement (e.g. number of number of students who got a B1-level certificate in language X at the end of the policy evaluated). Cost-effectiveness is computed through ratios. For example, the cost-effectiveness of different programmes aimed at improving the foreign language skills of a given group of adults can be measured in terms of ratio between the costs of a programme and the number of persons who passed the test certifying a level of knowledge equal to B1.

In policy evaluation, comparing programmes according to the criterion of efficiency means assessing which option displays the higher net benefit, that is, the higher difference between policy benefits and costs, where in the standard approach based on cost-benefit analysis (CBA, both benefits and costs are expressed in monetary terms. Benefits, in fact, are outcomes to which we can attach a monetary value.

Relevance, effectiveness, cost-effectiveness and efficiency are not the only possible evaluation criteria. A very important criterion is equity (or fairness) which may refer either to the distributive consequences of a policy for relevant groups of actors (e.g. defined according to their mother tongue).

7. Choosing methods and tools. In order to answer evaluation questions, appropriate methods and tools have to be chosen. A method is an ad hoc procedure designed for answering one or more evaluation questions. Methods, therefore, are “families of evaluation techniques and tools that fulfil different purposes” (European Commission 2008: 158), whereas a tool is defined as a standardised procedure having a particular evaluation function and that specifically operationalises a method. For example, assume that an evaluator

⁴ Some authors use the term “impact” for effects all the way at the end of the process, using “outcome” to describe an intermediate stage between “output” and “impact”. We shall, however, avoid the term “impact” and use the term “outcome” to denote effects in terms of the actual policy goal (for example, language vitality, language use, etc.).

has to collect data concerning the results of a policy which promotes a minority language in the schools of a given area. A method might be gathering the views of school directors, whereas a tool might be a survey through questionnaire or a focus group. The choice of methods and tools is also influenced by the financial resources that can be spent on the evaluation, timetable, and data available.

3.3 IMPLEMENTING EVALUATION

Implementing an evaluation can also be structured in five steps: (8) designing or assessing the programme theory, (9) designing or assessing programme implementation and process, (10) designing or examining the information system, (11) defining which data are needed to evaluate outputs and outcomes, or actually obtaining them, and (12) providing a final judgment and conclusions. Let us discuss these steps more closely.

8. Assessing programme theory. Programme theory is defined as the conceptualisation behind the policy, or, in other words, “the logic that connects its activities to the intended outcomes, and the rationale for why it does what it does” (Rossi et al. 2004: 44). The internal logic of a policy, therefore, is the plan linking all policy elements, that is, the public problem to tackle (or the needs to be satisfied), the policy objectives, the resources employed, the effects expected, etc.

In ex-ante evaluation, the analysis of the programme theory of policy alternatives is obviously a central part of the process of evaluation itself. In ex-post evaluation, examining the programme theory can be useful to determine whether the ex-ante evaluation (if any) has been conducted properly. The ex post evaluation of the programme theory can also serve to clarify whether objectives, official documents, etc. were sufficiently clear.

Evaluating the programme theory means, among other things, assessing whether the logical links between the policy elements are properly spelled out. The programme impact theory, therefore, is a causal theory, and it requires a clear description of causal sequence connecting policy measures with expected outcomes. Both the ex-ante and the ex-post assessment of the programme theory can be based on an analysis of policy documents, a reconstruction of decisions maker’ choices, but also resort to formal modelling.

The use of formal modelling is especially relevant in ex ante evaluation, as it can help to ground these cause-and-effect relationships in a theory of human action, and more specifically a theory of language choice. For example, the Policy-to-Outcome Path model, a.k.a. the P-TOP model (Grin 2003: 44-48), has been used as an analytical framework for the evaluation of the cost-effectiveness of several language policies aimed at supporting minority languages in Europe (Grin et al. 2002).

9. Assessing programme implementation and process. Evaluating the quality of policy implementation and management also implies the assessment of how organisational and budgetary resources have been used, or, alternatively, how it is planned to use them. The assessment of a programme implementation plan and process, therefore, can be regarded as an evaluation of how a language policy will be or has been managed.

10. Examining the information system. Realising an evaluation implies also examining the information system of a policy, that is, the “arrangements to store information on interventions, their context and progress (inputs, outputs and results) so that they can be accessed and inform decision makers, managers and evaluators” (European Commission 2008: 154). Indicators are the central element of any information system, since data and information are meaningful only if they are interpreted in the light of some indicators, which, in turn, make sense with respect to the evaluation questions and to programme theory. The assessment of the quality of an information system, therefore, includes not only the evaluation of the quality of indicators but also of the indicator system as a whole.

In ex-ante evaluation, the evaluator may contribute to both the design of a new information system or the assessment of the suitability of an existing system, whereas in ex-post evaluation, he assesses whether the information system as a whole has been appropriate in terms of its capacity to provide input for answering evaluation questions.

11. Obtaining data and evaluating outputs and outcomes. Data can be both quantitative and qualitative. As a general rule, “quantitative data are most likely to be used when aggregation and generalisation is required; and qualitative data when complexity and the finer details of experience need to be described” (European Commission 2008: 111). Indicators are usually populated through quantitative data, but in the field of multilingual

communication and LPP, these data are rarely available. For this reason, qualitative data collected through research tools such as focus groups and case studies are often necessary.

Measurement requires tools through which data are collected and analysed. Among others, individual stakeholder interviews, focus groups, case studies, and observational techniques are the most commonly used tools for collecting qualitative data, whereas social surveys are often used to collect quantitative data. The evaluator can use also secondary data and primary administrative source data routinely collected as part of policy management. Ethnographic observations are helpful when an evaluator wants to approach the insiders' views, in order to assess a policy "from within". Discourse analysis and conversation analysis are two forms of naturalistic inquiry complementing observational methods that may provide useful inputs for evaluation to the extent that they can be connected to some particular evaluation questions.⁵

When analysing the data collected, particular emphasis is placed on the estimation of net policy effects. The net effect is defined as the difference between the effect actually measured and the effect that would have appeared in absence of the policy. This presupposes the definition of a point of comparison, or counterfactual. The simplest case of counterfactual is the status quo or the absence of a policy.

A possible strategy for assessing the net effects of a policy is also to adopt experimental techniques, based on based on a comparison between a "treatment group" and a "control group". In many cases, however, for practical or ethical reasons, it is impossible to apply experimental techniques, and the evaluation of the net effects is necessarily based on other tools, such as econometric analysis. In ex-ante evaluation outputs and outcomes can only be estimated.

12. Providing a final judgment and conclusions. The final step of any evaluation is to provide a judgment and to draw up conclusions based on the evaluation criteria chosen. For example, in an ex-post evaluation, the analyst might focus on the effectiveness of the policy and on its distributive consequences. In an ex-ante evaluation, the evaluator could focus on the expected efficiency or

⁵ See for example Wodak and Krzyzanowski (2008) and Wodak and Meyer (2009) for an introduction to discourse analysis and Schegloff (2007) for conversation analysis. On their relevance to policy evaluation, see Grin, Sfreddo and Vaillancourt (2010), Chap. 2.

cost-effectiveness of alternative policy options. Cost-benefit analysis (CBA),⁶ cost-effectiveness analysis (CEA)⁷ and cost-utility analysis (CUA)⁸ are among the most commonly used tools for aggregating input indicators on the one hand, and outcome indicators on the other hand in order to compare them and come to a final judgment. CBA is used when both the advantages and drawbacks of a policy can be assessed in monetary terms, whereas CEA and CUA are preferred when advantages cannot be translated in monetary units.

Language policy analysis is necessarily partial, that is, it usually addresses one problem at time. Hence, it is not meant to solve language problems independently of local specificities. In addition, it is useful as a complement to decision making, not as a substitute to it. Language policy evaluation is necessarily based on a comparative approach, and therefore it is not supposed to identify the best solution, but only which solution among the options available is relatively better. Finally, notice that is a useful tool to assess incremental improvements rather than definitive solutions. As Weimer and Vining note, “policy problems rarely have perfect solutions, but some policies are better than others. A primary task for the policy analyst is to identify those policies that have the best prospects for improving social conditions, as measured in terms of specific goals and criteria” (2005: 209).

4. INDICATORS AND INDICATOR SYSTEMS

4.1 THE CONCEPT OF INDICATOR AND THE PROCESS OF INDICATOR DESIGN

An indicator may be defined as “the measurement of an objective to be met, a resource mobilised, an effect obtained, a gauge of quality or a context variable. An indicator produces quantified information with a view to helping actors concerned with public interventions to communicate, negotiate or make decisions. Within the framework of evaluation, the most important indicators are linked to the success criteria of public interventions” (European Commission 2008: 111).

⁶ Cost-benefit analysis is the object of a large literature. For an introduction see, among others, Brent (1997), Belli (2001) and Boardman et al. (2006). See in particular Mitchell and Carson (1989), Bateman et al. (2002) and Champ et al. (2003) for a discussion of indirect techniques used for estimating the monetary value of advantages.

⁷ See Levin and MacEwan (2001). See Grin (2003) and Grin et al. (2002) for an application of CEA to the promotion of regional or minority languages in Europe, and Gazzola (2006a, b) for a CEA of the language regime of the European Parliament.

⁸ See Brent (2003).

Indicators can measure facts but also opinions, and they can be derived from secondary sources such as censuses, or designed specifically for a single evaluation (ad hoc indicators). Indicators, therefore, provide information to decision makers during the analysis of a problem to solve, in the process of policy design and policy selection. They are also used to monitor the implementation of a policy and in ex-post evaluations. Indicators are also used to inform stakeholders, sponsors and the general public.

As a general rule, raw observations and quantitative information are not indicators per se. Indicators are theoretical constructs which must make sense with respect to the evaluative questions and programme theory that are populated by data. The nature and the role of an indicator, thus, are defined with respect to the needs of evaluators and not with respect to data.

Indicators are usually designed following two not mutually exclusive approaches. The first is based on a deductive strategy, whereas the second is inductive. The first approach is usually known as the lazarsfeldian paradigm—after the work of Paul Lazarsfeld (1958)—which is still the dominant paradigm in social sciences. The lazarsfeldian paradigm to indicator design is based on a deductive chain by which broad conceptual ideas are converted into instruments of empirical research. According to Lazarsfeld, the “process by which concepts are translated into empirical indices has four steps: an initial imagery of the concept, the specification of dimensions, the selection of observable indicators, and the combination of indicators into indices” (1958: 101).

The first step is based on the general construction of the concept, which is usually associated with the public problem the policy is supposed to address, for example, social exclusion, ethnic discrimination, language shift, etc. The second step consists of taking the original concept and specifying (or dividing) it in several operative dimensions, aspects, or components (see figure 3 below). Dimensions are “sometimes derived logically from the overall concept, or one aspect is deduced from another, or empirically observed correlations between them are reported. The concept is shown to consist of a complex combination of phenomena, rather than a simple and direct observable item” (Lazarsfeld 1958: 101).

Dimensions are usually decomposed in several sub-dimensions, and this process goes on until these dimensions and sub-dimensions get sufficiently

clear and precise to be translated into individual indicators. An indicator does not need to be further decomposed and it is specific enough to be populated by data. In the Lazarsfeldian paradigm, therefore, indicators are tools that semantically indicate a concept, or, more precisely, its constitutive dimensions, and they are the final link between the concept and the techniques used to collect observations. If necessary, indicators can be aggregated into indices. Indices aim at giving a synthetic result usually based on a weighted sum of indicators.

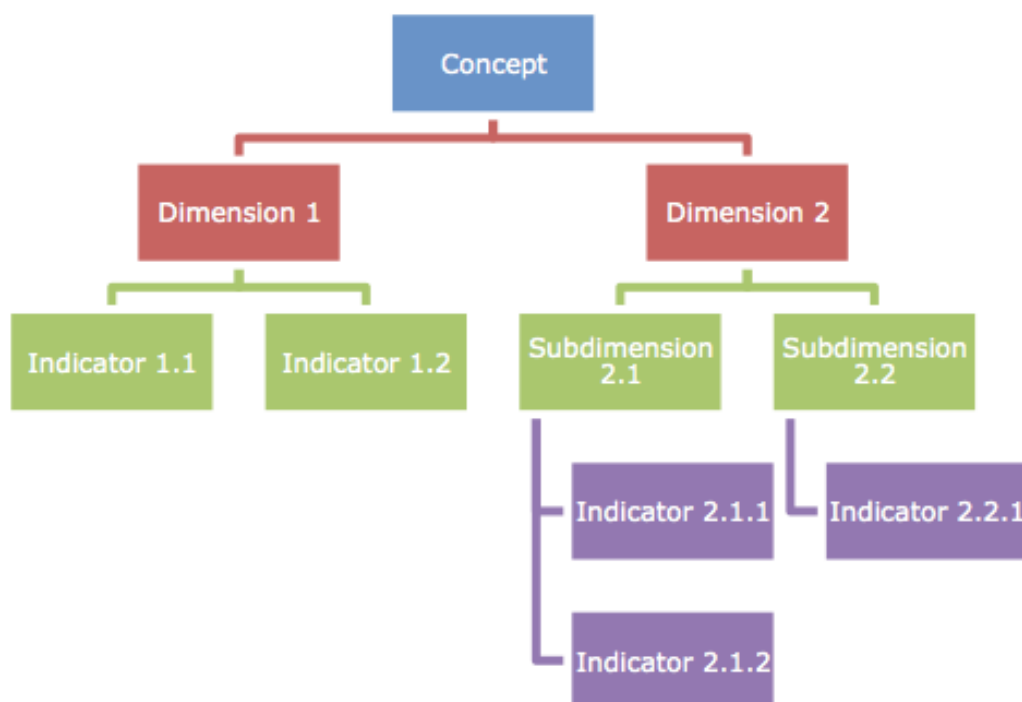


Figure 3 : A simple example of decomposition of a concept into indicators

Several conceptual and practical difficulties arise in the process of indicator design. As noted by Lazarsfeld, “the fact that each indicator has not an absolute but only a probability relation to our underlying concept requires us to consider a great many possible indicators” (1958: 103). In other words, it is usually not possible to find a one-to-one relationship between an indicator and the concept that the indicator is assumed to indicate. Clearly, an appropriate relationship with the underlying concept is necessary to fulfil the “validity” criterion (see section 4.2 below), but often the same indicators can indicate different things, and different indicators can indicate the same concept.

The second possible strategy for indicator design is inductive. Starting from the data available, the analyst tries to figure out how they can be elaborated and transformed into indicators suitable to operationalising the concept. Usually the deductive and the inductive approaches are used simultaneously, since the

specification of the concepts is developed in parallel with the analysis of the data already available. The design of the information system of a policy is developed in parallel with the process of policy making, and therefore the decomposition of concepts usually reflects the process through which general objectives are articulated in specific objectives (cf. table 2 below).

Since evaluation must be useful to stakeholders and decision makers, policy indicators should not be designed following an exclusively top-down process, that is, a process during which the evaluator designs, at the same time, both the policy and the indicators system. Indicators should be also discussed and negotiated in a participative way with the stakeholders, since the process of indicator design in policy evaluation is also a process through which policy objectives are re-defined and interpreted in the light of stakeholders' points of view.⁹

4.2 TYPOLOGIES AND DESIRABLE PROPERTIES OF INDICATORS

Indicator can be classified according to different typologies. In relation with the processing of information, we shall make a distinction between elementary, derived, and compound indicators. Elementary indicators are based on very simple dimensions, such as "the number of speakers of language X". Derived indicators (or ratios) are constructed on the basis of (usually) two elementary indicators, whereas compound indicators (or indices) are defined as the weighed sum of several elementary or derived indicators.

In relation with the scope of information, a distinction must be made between context and policy indicators. Context indicators are used to characterise the situation of a given population of a given territory. They are usually applied to the population and territory as a whole, and not only to the specific target territory or population of the policy. Context indicators are used to assess the relevance of a policy. By contrast, policy indicators are applied only to that particular part of the population and/or territory concerned by the policy. These indicators are used to assess the effects of a policy and therefore whether it has been successful.

In relation with the phases of completion of the policy, we shall distinguish between input (resource), output, and outcome (result) indicators (cf. table 1

⁹ This procedure is based on a "bottom-up" approach harking back to constructivism, which is usually opposed "neopositivism". On this point, see European Commission (2008), Palumbo (2003) and Stame (2001).

above for definitions). Providing accurate description of actual processes through ethnographic research is only a preliminary step towards the design of indicators, since enumerating the consequences entailed by a language policy on a specific population in a given context at a given time is not necessarily equivalent to providing adequate indicators of the effects of a policy (cf. the "adequacy" criterion below). In relation with evaluation criteria, a distinction can be made, among others, between relevance, effectiveness, efficiency, cost-effectiveness, and fairness indicators. Other indicators can be designed according to the evaluative criteria chosen.

The four typologies of indicators presented should not be considered as mutually exclusive, as the same indicator can be at the same time a programme indicator, an outcome indicator and an effectiveness indicator. What changes is its role with respect to the goal of the analysis.

Let us now turn to the assessment of the quality of indicators and indicator systems. This phase is an important step in the evaluation process (cf. § 3.3 above). Even if no clear consensus in the literature exists, several criteria are used to define an indicator as "good":¹⁰

1. Validity (or normativity). An indicator should avoid ambiguities, and therefore the correspondence should be as clear as possible between the indicator and the object (or the concept) it is deemed to reflect (e.g. an outcome).
2. Reliability, which refers to the fact that if two different people under identical conditions take the same measurement, the result, within a certain margin of error, should be an identical indicator value.
3. Sensitivity. A policy indicator has to vary significantly when the policy is implemented and displays its effects. Sensitivity implies that indicators should not be too general and that their variation (positive or negative) should be univocal.
4. Stability. An indicator should be quantifiable on a regular basis so that the evaluation can generate time series.
5. Relevance and Adequacy. Indicators have to be relevant with respect to the objective pursued and meet the cognitive needs of the researchers.

¹⁰ We rely mostly on European Commission (2008: 118-119), Palumbo (2001: 346-347), Fitz-Gibbon and Tymms (2002) and Perret (2002: 27). See also Solé i Camardons and Torrijos (2008: 93) and Johnson and Doucet (2006: 46).

6. Feasibility. It must be possible to populate the indicators through primary or secondary data.
7. Representativeness. The more an indicator is able to summarise in a single value a large amount of phenomena, the more useful it is.
8. Intelligibility. A good indicator has to be understandable by everyone involved in the policy.
9. Timeliness of the information used to populate the indicator. In certain cases it can take a long period of time before the information can really be used for evaluation. The criterion of timeliness has to be taken into account especially for those indicators related to policies aimed at having effects in the long run. This is usually the case for language policies, whose effects can be assessed only after a relatively long period of time, e.g. one, five, or ten years.
10. Comparability. An indicator is useful if it allows comparisons between different policy measures and between different policies.
11. Power, that is, its capacity of distinguishing the various aspects of a phenomenon (e.g. distinguish between the effects of short term and long term training).

Let us now turn to the desirable qualities of an indicator system. We should distinguish between:

1. Coverage: the system of indicators should cover the majority of the policy measures.
2. Balance: there should be a good balance between different types of indicators (e.g. input, output and outcome indicators).
3. Selectivity: the system of indicators should be simple, that is, it should not include too many indicators.
4. Relevance: a system of indicators is relevant if it includes those indicators that are related to measures or themes with significant implications in terms of decision-making (that is, if it includes relevant indicators).

A further issue related to indicators and indicator systems concerns the use of indicators. It is important to be aware that indicators are not simply convenient tools for representing a reality which exists independently from the indicators themselves. The use of indicators for evaluation can have a direct

impact on actors' behaviour, since they could do something that they would not have done otherwise in order to cause a change in the value of an indicator towards a desired direction. In some cases, the change in actors' behaviour is precisely the goal of the policy, and indicators can provide an appropriate incentive for actors. However, in some cases the use of indicators can also have adverse effects.¹¹ For example, the indicator system can cause unanticipated consequences if the system causes the operators to work for the indicator rather than for the result (European Commission 2008: 123). This effect is sometimes named "goal displacement" (Perrin 1998). A well-known example in education is "teaching to the test".

5. EXAMPLES OF EXISTING LINGUISTIC INDICATORS

Developing indicators for the evaluation of language policies still is not a very common practice.¹² Nevertheless, some very interesting and promising experiences in language policy indicator design exist, in particular in multilingual countries such as Canada or Spain, and international organisations like the EU or UNESCO that have a long-lasting tradition in linguistic diversity management. Let us discuss some examples in light of the preceding sections.

The Office québécois de la langue française (OQLF — Quebec Board of the French language) regularly publishes a set of reports containing a large quantity and variety of demolinguistic indicators regarding the vitality of languages in Quebec (Castonguay 2005), as well as several indicators related to the use of languages in different domains, such as such public administration, cultural industry, education (Bouchard 2008) or economic activity (Bouchard 2006). These data are used to monitor the situation of languages in Quebec—and that of French in particular—and they are collected within the framework of Quebec's language policy.¹³ Studies concerning the evaluation of the vitality of communities speaking minority languages in Canada are published by the Office of the Commissioner of Official languages (Johnson and Doucet 2006).¹⁴

¹¹ Cf. European Commission (2008: 123-124) for a more extensive discussion on adverse effects of indicators in socio-economic programmes.

¹² In this section we focus on linguistic indicators provided by public authorities. However, literature in LPP is also a source for language policy indicators. Grin (2003: 105-106), for example, proposes a series of indicators in six areas of intervention in favour of minority languages mentioned in the European Charter for Regional or Minority Languages (RMLs).

¹³ On the Quebec language policy, and in particular on the Charter of the French Language, see, among others, Edwards (1998) and Corbeil (2007).

¹⁴ Notice, however, that the study of Johnson and Doucet (2006) focuses on the vitality of communities in general and not specifically on languages. Languages are regarded as an element of minorities' vitality.

The goal of indicators published by the OQLF is to provide a picture of the sociolinguistic situation of Quebec, in particular as regards the use of languages, and to keep track of changes over time. As regards languages used in the job market and on the workplace, the OQLF publishes several complex indicators monitoring the evolution of unemployment rate and the education of the workforce according to the first language of speakers. Examples include "frequency of use of French and English on the workplace by geographical region", or "main working language by sector, mother tongue and region"(Bouchard 2006).

Indicators related to education (from primary school to university) are context indicators whose goal is to present a faithful account of the situation as regards the language in education and monitor its evolution (Bouchard 2008: 99). Examples of such indicators are the "number of pupils eligible for teaching in English in Quebec" (indicator 4.5), or "enrolments at the University according to the mother tongue of students" (indicator 4.19). These indicators are elementary or derived context indicators aimed at providing a picture of an existing situation. They could be used as policy indicators if linked to a specific policy to be evaluated.

An interesting indicator developed by the OQLF is the index of linguistic vitality or index of continuity (Castonguay 2005: 11). The index of linguistic vitality for language X in a given domain (e.g. family) is defined as the ratio between the number of people (native and non native speakers) who use X in that domain and the number of people whose mother tongue is X. A language used by its native speakers and also by other people as L2, will display an index larger than one. On the contrary, a ratio smaller than one will denote a low vitality, typically because people speaking X as a mother prefer to use another language. The index of linguistic vitality, therefore, indicates two dimensions of the concept of "linguistic vitality", namely, persistence and attractiveness. Persistence refers to the fact that the people speaking language X as mother tongue use it in different domains. Attractiveness refers to capacity of a given language to be used in a given domain also by non native speakers. Confronting the value of the index across time, therefore, should inform about the persistence of use of language X among native speakers, and about the evolution of attractiveness among non native speakers.

The issue of linguistic vitality has also been addressed by the United Nations Educational, Scientific and Cultural Organization (UNESCO). In 2003 UNESCO published a report providing a list of six "major evaluative factors of language

vitality” (UNESCO 2003), namely, (1) intergenerational language transmission; (2) absolute number of speakers; (3) proportion of speakers within the total population; (4) trends in existing language domains; (5) response to new domains and media; and (6) materials for language education and literacy. These “factors” can be regarded as specifications (or dimensions) of the concept of linguistic vitality (cf. section 4.1 above). Dimensions (2) and (3) may be directly linked to quantifiable indicators in a straightforward way, but UNESCO does not provide a set of quantifiable indicators related to the four other factors. However, it provides qualitative and descriptive scales to rank language vitality according to the six dimensions.

Let us now turn to Spain. Both the Basque and the Catalan regional authorities collect different linguistic indicators for monitoring the sociolinguistic situation of the Basque and the Catalan language respectively.¹⁵

The Hizkuntza Politikarako Sailburuordetza (“The Sub-ministry for Language Policy” of the Basque Government) has set up the Euskal Herriko Hizkuntza-adierazleen Sistema—EAS (“System of linguistic indicators of Euskal Herria”), which is “an instrument designed to provide local government, agents and organisations engaged in the standardisation of the use of Basque, as well as the inhabitants of this region, with detailed information about the status of the Basque language and its development within the Basque country as a whole”.¹⁶

The EAS indicators are divided in five groups, that is, context, resources, legislation, sectoral programmes and results. These names, however, should not be confused with the terminology adopted in section 4.2, since the two groups of terms do not perfectly overlap. In the EAS, for example, result indicators include indicators like “level of language competence (> = 16), according to age”, which is not in itself an indicator of result or outcome. Depending on the analytical perspective adopted, this indicator can be a context or a policy indicator. It can also be an output indicator, as opposed to result/outcome indicators, if we regard the outcome of a language policy in favour of a minority language as the time per day spent in speaking Basque (rather than Castilian, for example). Resource indicators refer to general budgetary expenditures of public authorities for the promotion of Basque divided by sector (teaching, linguistic landscape and signs, etc.). These data can be used to populate input indicators for specific policy evaluations.

¹⁵ On language policies in Spain, see for example the review in Castillo Lluich and Kabatek (2006).

¹⁶ Cf. http://www1.euskadi.net/euskara_adierazleak/about.apl?hizk=i

EAS indicators provide general information on the use of Basque in family, its intergenerational transmission, the level of competence in Basque, and attitudes towards policies for the promotion of Basque. In addition, the system provides different sociolinguistic indicators or simply information on several domains like education (e.g. "number of doctoral thesis published in different languages at universities in the Basque Country"), public administration (e.g. "degree of achievement of language profiles in public administration"), the workplace (e.g. "Basque language use in enterprises with more than 100 staff"), mass media (e.g. "journals and magazines in Basque") and Basque language learning for adults (e.g. "number of students by sex, education, profession").

One of the richest sets of linguistic indicators available is that of the Generalitat de Catalunya ("Government of Catalonia").¹⁷ For reasons of space, it is not possible to discuss these indicators in detail. Suffice it to say that, among other things, the Government of Catalonia publishes the *Sistema d'Indicadors Lingüístics a Catalunya—SIL—*("Catalan system of linguistic indicators") a series of more than eighty linguistic indicators related to seven domains of language use, namely: population, public administration, the educational system, mass media and cultural industries, the socioeconomic world, the world of associations and health institutions (Solé Camardons and Torrijos 2008, Secretaria de Política Lingüística 2009). These seven social domains are articulated in more than thirty sub-domains. The collection of these indicators is part of the general *Pla estadístic de Catalunya 2006-2009* ("Statistical Plan of Catalonia"). In addition, it also publishes different reports on the use of Catalan in various sociolinguistic domains and other statistical data related to the demolinguistic situation of Catalan.

The goal of the SIL is twofold: "a) understanding and processing data on language knowledge, availability and use in the general population and in specific sectors of society; b) understanding and processing data on the results and on the sociolinguistic impact of language policy actions" (Solé Camardons and Torrijos 2008: 92). Hence, the SIL is not only a cognitive but, at least in principle, also an explicit evaluation instrument. SIL indicators are organised in four groups, namely, competence ("competència"), use ("usos"), supply ("oferta lingüística"), and consumption and audience ("consum i audiència").

¹⁷ In the Catalan context, a linguistic indicator is defined as "a serial variable or the relationship between quantitative variables, the extent of which over time provides important information about the linguistic situation of an organisation, of its environment or of society in general, and helps with making policy decisions and planning techniques" (Solé Camardons and Torrijos 2008: 92). For a more in-depth discussion of linguistic indicators and evaluation in the Catalan case, see Jaime et al. 1995, Simó et al. 2006; Solé Camardons 2003, Solé Camardons and Torrijos 2008; Mur i Petit 1999, Fabà et al. 2000.

The SIL follows a lazarsfeldian paradigm, in which indicators are derived from a general concept. Some of the main criticisms to SIL are related to (i) the lack of international comparability, (ii) the lack of interest of citizens for these indicators, (iii) and, more importantly, the absence of cost, effectiveness and cost-effectiveness indicators of language policy (Solé Camardons and Torrijos 2008: 101-102).

In terms of the typology presented in section 4.2, this implies that the “comparability”, “intelligibility” and “power” of indicators should be improved. The lack of cost, effectiveness and cost-effectiveness indicators shows that SIL still does not sufficiently emphasise the role of indicators in relation to evaluation. As regards the system as a whole, Solé i Camardons and Torrijos note that the large number of indicators used could lead to contradictory conclusions (2008: 100). A possible strategy is precisely to increase the “selectivity” and “relevance” of the system as a whole (cf. section 4.2 above).

The SIL, however, is also interesting as part of a larger set of indicators called the Indicators quantitius sectorials (“Quantitative sector indicators”), which also include other data (Secretaria de Política Lingüística 2009). Non-SIL indicators provide useful information on people’s attitudes vis-à-vis language policy. For example, the indicators collected by the Catalan Consumer Agency include the “number of language complaints lodged in 2008” or the “number of complaints received about the lack of spoken service in Catalan in 2008”. This constitutes useful information on the effectiveness of communication in business (where effectiveness is interpreted as the capacity of business to persuade, influence or charm customers).¹⁸

At the supranational level, the European Commission publishes several data focusing on use and knowledge of foreign languages in Europe. These data have been published in 2001, 2005 and 2006 as a result of three large Eurobarometer surveys.¹⁹ The last Eurobarometer survey, for example, contains demographic data regarding European citizens’ mother tongue and their skills in foreign languages (European Commission 2006). In addition, it provides data regarding Europeans’ attitudes towards different foreign

¹⁸ Linguistic indicators measuring customer satisfaction are also produced in Canada. Augen (2005, quoted in Johnson and Doucet 2006: 49), for example, designed several indicators to assess the rate of satisfaction of public services offered in the two Canadian official languages, focusing on two dimensions, namely, accessibility and performance. These indicators can be used as part of the evaluation of the effectiveness of communication in the public sector, where effectiveness in this case is interpreted as the capacity of the institution to provide inclusive and shared communication towards citizens. On this point, cf. Gazzola and Grin (2007) who present a general framework for the evaluation of the effectiveness and cost-effectiveness of institutional communication in multilingual contexts.

¹⁹ Cf. http://ec.europa.eu/public_opinion/archives/eb_special_en.htm.

languages,²⁰ data regarding ways of learning and the frequency of use of foreign languages in daily life. These data have been collected within the framework of the language policy adopted by the EU, which recommends the learning of at least two foreign languages since from early age.²¹

So far data collection has not been followed by any specific elaboration of language policy indicators. A partial exception is the European Indicator of Language Competence, whose design was launched by the European Commission in 2005 (European Commission 2005). The technical preparation work started in 2007 and the tests should be carried out in 2010. This indicator aims to provide Member States, policy makers, teachers and practitioners with reliable comparative data on foreign language competence across the European Union. It will provide knowledge about the multilingual capacities of young people, on where good practice and performance can be found, and on progress towards the objective of improving foreign language learning.

Very little is known about evaluation of language regimes in international institutions, both as regards internal and external communication. A language regime is a form of public policy determining the set of official and working languages and the rule concerning the use of these languages for institution's communication. Evaluating language regimes implies, among other things, the development of indicators for assessing the "performance" (process) and the "results" (product) of translation and interpreting services (TIS). The EU, the largest user of TIS worldwide, still does not have a system of indicators to track the performance of its TIS and, more generally, to evaluate the effects of its language regime.

As regards translation, a report of the European Court of Auditors concluded that

none of the translation services [of the institutions of the EU] has a tool for measuring client satisfaction or procedures for complaints about quality. Only the Commission has guidelines for quality control. Quality indicators, such as the number of errors found per page, are available at the Parliament but not at the Commission or the Council. [And] with the exception of the Commission for 2002, none of the institutions had calculated their total translation cost or the average cost per page translated" (European Court of Auditors 2006: 4).

²⁰ For example, respondents were asked to express their opinion as to the "utility" of different foreign languages.

²¹ Cf. European Council, Barcelona, 15-16 March 2002, Presidency Conclusions, part I, 43.1.

The audit report concludes that “consistent monitoring through the use of performance indicators and procedures for ensuring adequate management information should be implemented” (European Court of Auditors 2006: 5).²²

Beside a systematic collection of data regarding costs of TIS, the European court of Auditors (2006: 17) suggests designing indicators and collecting statistics on (a) the actual time spent, versus the standard or estimated time, on the translation of a specific document; (b) the number of pages revised and the time spent on revision; (c) the production and productivity of individual translators; (d) the volume of translations outsourced automatically by the planning units and those outsourced by the translation units; (e) failure to observe delivery deadlines; (f) the impact of IT translation tools on productivity; (g) crossed data, such as the correlation between workload and outsourcing.

Note that there is a conceptual difference between the evaluation of the performance of TIS in itself on the one hand, and the evaluation of TIS as an element of a broader language policy on the other hand. The first analytical perspective is strictly internal and it is closed to management sciences, whereas the second approach is consistent with public policy analysis. Obviously, the two approaches are not mutually exclusive and some data and indicator may be commonly used.

The development of policy indicators based on the Court suggestions is a promising research object for LPP. More specifically, indicators of performance could be put into relation with a broader system policy indicators of effectiveness and/or fairness. In certain cases, this relation can be relatively straight forward, for example, performance indicators such as the number of errors per page.

In other cases, however, this relationship may be more indirect. Research has shown that misunderstandings in multilingual communication settings are attributed to different factors, among others, (i) the lack of foreign language skills (both passive and active) of political representatives (e.g. MEPs) and/or civil servants, and (ii) the mental fatigue caused by working in a foreign language.²³ Performance indicators such as the impact of IT translation tools on productivity or TIS (cf. point f above), for example, could contribute to

²² For interpreting, see European Court of Auditors (2005).

²³ Cf. Grin and Gazzola (2009).

effectiveness in communication if an empirical link between TIS and a significant reduction of multilingual civil servants' mental fatigue can be established.

It would also be advisable to design indicators reflecting the effectiveness of language regime in meeting users' expectations, for example, the number and type of complaints related to communication lodged by civil servants (internal communication) or citizens (external communication), etc.

It is worth repeating that "indicator" is not synonymous with data. In practical terms, an elementary indicator can coincide with data (e.g. number of speakers of language X). However, strictly speaking, indicators should be referred to some explicit evaluative or research questions and programme theory, whereas data are simply the input used to populate indicators.

CONCLUDING REMARKS

Evaluation theory is usually applied to deliberate public policies aimed at specific results, such as pollution reduction, increasing the vitality of a minority language, etc. However, in some cases, no explicit or deliberate language policy is adopted by public institutions and what the analyst can observe are simply practices. For example, sometimes there is no explicit language policy regarding the official use of language(s) in a country. Generally speaking, since institutions cannot abstain from using language, there is always at least one default language policy. Furthermore, as regards policy timing, policy outcomes are normally assessed as soon as the programme has been completed or at the end of the exposure of the target population to it. Some language policies, however, are not necessarily designed in such a way that we can define an "end". For example, an international institution chooses a set of official and working languages to communicate, rather than to achieve a given goal within a certain period of time.

How can we deal with this problem? Even without an explicit language policy, we can assess observed situations as if each were the result of a deliberate way of handling multilingual communication, and then compare these situations through indicators. In other words, policy evaluation theory is used as a framework for structuring observation and studying communication situations in the light of counterfactual analysis, that is, the study of selected aspects of a policy with respect to a given term of comparison.

Hence, even if language policies are sometimes not implemented in order to achieve a specific result, this does not imply that they cannot be evaluated. First, the language policies adopted in one context can be compared synchronically to a similar policy adopted elsewhere. In addition, language policies can be compared diachronically. For example, the EU's language regime before and after the increase in the number of official and working languages due to the 2004 enlargement, or the language regime of the United Nations before and after the inclusion of Arabic as a new official language in 1973, etc.

Evaluation can make a valuable contribution to LPP, not only because all public policies should be the object of careful analysis, but also because language policy evaluation is a very promising domain of research since "the contribution of evaluation is potentially greatest in innovative policy areas where achieving success cannot be taken for granted and where implementation is not always straightforward" (European Commission 2008).

REFERENCES

- Aunger, Edmund A. (2005) "Diversité régionale et inégalité politique : les minorités de langue officielle et le problème de deux poids deux mesures", in Floch, William and Yves Frenette (eds.) *Vitalité des communautés, confiance des communautés*. Forum de recherche sur les langues officielles. Ottawa: Patrimoine canadien.
- Bateman, Ian, Richard T. Carson, and Brett Day (2002) *Economic Valuation with Stated Preference Techniques*. Cheltenham: Edward Elgar.
- Belli, Pedro (2001) *Economic Analysis of Investment Operations: Analytical Tools and Practical Applications*. Washington (DC): World Bank Publications.
- Boardman, Anthony E., David H. Greenberg, Aidan R. Vining, and David L. Weimer (2006) *Cost-Benefit Analysis. Concepts and Practice* (3rd ed.). Upper Saddle River (NJ): Prentice Hall.
- Bouchard, Pierre (2006) *Langue du travail : indicateurs relatifs à l'évolution de la population active et à l'utilisation des langues au travail en 2001*. Québec: Office québécois de la langue française.
- Bouchard, Pierre (2008) *La langue de l'enseignement : indicateurs pour l'éducation préscolaire, l'enseignement primaire et secondaire, le collégial et l'université*. Québec: Office québécois de la langue française.

- Brent, Robert J. (1997) *Applied Cost-Benefit Analysis*. Cheltenham: Edward Elgar.
- Brent, Robert J. (2003) *Cost-Benefit Analysis and Health Care Evaluations*. Cheltenham: Edward Elgar Publishing.
- Castillo Lluch, Mónica and Johannes Kabatex (eds.) (2006) *Las lenguas de España. Política lingüística, sociología del lenguaje e ideología desde la Transición hasta la actualidad*. Madrid: Iberoamericana.
- Castonguay, Charles (2005) *Les indicateurs généraux de vitalité des langues au Québec : comparabilité et tendances 1971-2001*. Québec: Office québécois de la langue française.
- Champ, Patricia A., Kevin J. Boyle, and Thomas C. Brown (eds.) (2003) *A Primer on Nonmarket Valuation*. Dordrecht: Kluwer Academic Publishers.
- Corbeil, Jean-Claude (2007) *L'embaras des langues. Origine, conception et évolution de la politique linguistique québécoise*. Montreal: Québec Amérique.
- Edwards, John (ed.) (1998) *Language in Canada*. Cambridge (UK): Cambridge University Press.
- European Commission (2005) *The European Indicator of Language Competence (Communication from the Commission to the European Parliament and the Council), COM(2005) 356 final*. Brussels: European Commission.
- European Commission (2006) *Europeans and their Languages, Special Eurobarometer 243*. Brussels: European Commission.
- European Commission (2008) *EVALSED: The resource for the evaluation of socio-economic development*. Luxembourg: Office for Official Publications of the European Communities.
- European Court of Auditors (2005) *Special report No 5/2005 Interpretation expenditure incurred by the Parliament, the Commission and the Council, together with the institutions' replies*, Official Journal of the European Union, C 291/1 (23.11.2005). Brussels: European Court of Auditors.
- European Court of Auditors (2006) *Special Report No 9/2006 concerning translation expenditure incurred by the Commission, the Parliament and the Council together with the institutions' replies*, Official Journal of the

European Union, C 284/01 (21.11.2006). Brussels: European Court of Auditors.

Fabà, Albert, Pilar López, Joan Solé Carnardons, and Noemí Ubach (2000) "Ofercat: indicadors sobre l'oferta de català a Santa Coloma de Gramenet", *Llengua i Ús - Revista tècnica de Política Lingüística*, (18), pp. 53-58.

Fitz-Gibbon, Carol Taylor and Peter Tymms (2002) "Technical and ethical issues in indicator systems: Doing things right and doing wrong things", *Education Policy Analysis Archives*, 10 (6), pp. 1-28.

Gazzola, Michele (2006a) "La gestione del multilinguismo nell'Unione europea", in Carli, Augusto (ed.) *Le sfide della politica linguistica di oggi : fra la valorizzazione del multilingualismo migratorio locale e le istanze del plurilinguismo europeo*, pp. 17-117. Milan: FrancoAngeli.

Gazzola, Michele (2006b) "Managing Multilingualism in the European Union: Language Policy Evaluation for the European Parliament", *Language Policy*, 5 (4), pp. 393-417.

Gazzola, Michele and François Grin (2007) "Assessing efficiency and fairness in multilingual communication: Towards a general analytical framework", *AILA [Association Internationale de Linguistique Appliquée] Review*, 20, pp. 87-105.

Geva-May, Iris and Leslie A. Pal (1999) "Good Fences Make Good Neighbours: Policy Evaluation and Policy Analysis — Exploring the Differences", *Evaluation*, 5 (3), pp. 259-277.

Grin, François (2003) *Language Policy Evaluation and the European Charter for Regional or Minority Languages*. Basingstoke: Palgrave Macmillan.

Grin, François, Tom Moring, Durk Gorter, Johan Häggman, Dónall Ó Riagáin, and Miguel Strubell (2002) *Support for Minority Languages in Europe*, European Centre for Minority Issues and European Bureau for Lesser Used Languages
<http://europa.eu.int/comm/education/policies/lang/languages/langmin/files/support.pdf>.

Grin, François and Michele Gazzola (2009), "Efficiency and Fairness - Deliverable No. 3", *DYLAN Working Paper N° 3*, University of Geneva

- Jaime, Laura, Xavier Mitjà, and Joan Solé Camardons (1995) "Objectius i indicadors de resultat: exemples de formulació", *Llengua i Ús - Revista tècnica de Política Lingüística*, (2), pp. 15-20.
- Jernudd, Björn H. (1971) "Notes on economic analysis for solving language problems", in Rubin, Joan and Björn H. Jernudd (eds.) *Can Language Be Planned?*, pp. 263-276. Honolulu: The University Press of Hawaii.
- Johnson, Marc L. and Paule Doucet (2006) *A sharper view: Evaluating the vitality of official language communities*. Ottawa: Office of the Commissioner of Official Languages of Canada.
- Knoepfel, Peter, Corinne Larrue, Frédéric Varone, and Michael Hill (2007) *Public Policy Analysis*. Bristol: The Policy Press. [Original: *Analyse et pilotage des politiques publiques* (2nd ed.), Geneva-Basel: Helbing & Lichtenhahn, 2006.].
- Lazarsfeld, Paul F. (1958) "Evidence and inference in social research", *Daedalus*, 87 (4), pp. 89-109.
- Levin, Henry M. and Patrick J. MacEwan (2001) *Cost-Effectiveness Analysis : Methods and Applications* (2nd ed.). Thousand Oaks London: Sage.
- Mitchell, Robert C. and Richard T. Carson (1989) *Using Surveys to Value Public Goods : the Contingent Valuation Method*. Washington (DC): Resources for the Future Press.
- Mur i Petit, Rosa (1999) "Indicadors d'ús del català en la documentació local 1998", *Llengua i Ús - Revista tècnica de Política Lingüística*, (16), pp. 37-45.
- Palumbo, Mauro (2001) "Valutazione di processo e d'impatto: l'uso degli indicatori tra meccanismi ed effetti", in Stame, Nicoletta (ed.) *Valutazione 2001. Lo sviluppo della valutazione in Italia*, pp. 328-360. Milan: FrancoAngeli.
- Palumbo, Mauro (2003) "Gli indicatori valutativi", *Rassegna italiana di valutazione*, 7 (27), pp. 107-129.
- Perret, Bernard (2002) "Indicateurs sociaux. État de lieux et perspectives" *Les Papiers du CERC 2002-01*. Paris: Conseil de l'emploi, des revenus et de la cohésion sociale.
- Perrin, Burt (1998) "Effective Use and Misuse of Performance Measurement", *The American Journal of Evaluation*, 19 (3), pp. 367-379.

- Ricento, Thomas (2006) "Language Policy: Theory and Practice - an Introduction", *An Introduction to Language Policy. Theory and Method*, pp. 10-23. Oxford: Blackwell.
- Rossi, Peter H., Mark W. Lipsey, and Howard E. Freeman (2004) *Evaluation: A Systematic Approach* (7th ed.). Thousand Oaks, London: Sage.
- Rubin, Joan (1971) "Evaluation and Language Planning", in Rubin, Joan and Björn H. Jernudd (eds.) *Can language be planned? Sociolinguistic Theory and Practice for Developing Nations*, pp. 217-252. Honolulu: The University Press of Hawaii.
- Schegloff, Emanuel A. (2007) *Sequence Organization in Interaction. A Primer in Conversation Analysis*. Cambridge (UK): Cambridge University Press.
- Secretaria de Política Lingüística (2009) *Informe de política lingüística 2008 - VII Indicadors quantitius sectorials*. Barcelona: Generalitat de Catalunya.
- Simó, Montse, Joan Solé i Comardons, and Anna Torrijos (2006) "Aproximació a l'avaluació de la campanya «Dóna corda al català»", *Llengua i Ús - Revista tècnica de Política Lingüística*, (35), pp. 26-32.
- Solé Camardons, Joan (2003) "El Sistema d'Indicadors Lingüístics (SIL): finalitat i característiques generals", *Noves SL. Revista de Sociolingüística*, 4 (2), pp. 1-18.
- Solé Camardons, Joan and Anna Torrijos (2008) "El sistema d'indicadors lingüístics a Catalunya", *Llengua i Ús - Revista tècnica de Política Lingüística*, (43), pp. 92-103.
- Stame, Nicoletta (2001) "Tre approcci principali alla valutazione: distinguere e combinare", in Palumbo, Mauro (ed.) *Il processo di valutazione. Decidere, programmare, valutare*, pp. 21-46. Milan: FrancoAngeli.
- Thorburn, Thomas (1971) "Cost-Benefit Analysis in Language Planning", in Rubin, Joan and Björn H. Jernudd (eds.) *Can Language Be Planned?*, pp. 253-262. Honolulu: The University Press of Hawaii.
- UNESCO (2003) *Language vitality and endangerment*, UNESCO ad hoc expert group on endangered languages. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO).
- Vaillancourt, François (1983) "The economics of language and language planning", *Language Problems & Language Planning*, 2 (7), pp. 162-178.

Vaillancourt, François (1985) "Les écrits en économie de la langue : brève revue et introduction au recueil", in Vaillancourt, François (ed.) *Économie et langues*, pp. 11-25. Québec: Conseil de la langue française.

Weimer, David L. and Aidan R. Vining (2005) *Policy Analysis: Concepts and Practice* (4th ed.). Englewood Cliffs: Prentice-Hall.

Wodak, Ruth and Michal Krzyzanowski (eds.) (2008) *Qualitative Discourse Analysis in the Social Sciences*. Basingstoke: Palgrave Macmillan.

Wodak, Ruth and Michael Meyer (eds.) (2009) *Methods for Critical Discourse Analysis* (Introducing Qualitative Methods series). London: Sage Publications.

World Bank (2004) *Monitoring & Evaluation: Some Tools, Methods and Approaches*. Washington (DC): The World Bank.