Epistemological and Theoretical Foundations in Language Policy and Planning

Michele Gazzola
Federico Gobbo
David Cassels Johnson
Jorge Antonio Leoni de León

Epistemological and Theoretical Foundations in Language Policy and Planning

Michele Gazzola · Federico Gobbo ·
David Cassels Johnson ·
Jorge Antonio Leoni de León

Epistemological and Theoretical Foundations in Language Policy and Planning

Michele Gazzola
School of Applied Social & Policy
Sciences
Ulster University
Belfast, UK

Federico Gobbo
Faculteit der Geesteswetenschappen
University of Amsterdam
Amsterdam, Noord-Holland,
The Netherlands

David Cassels Johnson
College of Education
University of Iowa
Iowa City, IA, USA

Jorge Antonio Leoni de León
Facultad de Letras
University of Costa Rica
San José, Costa Rica

Cover illustration: © John Rawsterne/patternhead.com

CONTENTS

LIST OF FIGURES

Epistemological and Theoretical Foundations in Language Policy and Planning: Introduction

Abstract This introductory chapter reviews the historical development of the field of Language Policy and Planning (LPP), outlines the structure of the book, and introduces conceptual concerns and challenges that circulate throughout the chapters. The authors review the foundations of the field in classic language planning theory and examine the evolution of disciplinarily and interdisciplinary approaches. Chapters in the book examine (socio)linguistic foundations, critical empirical research, the public policy approach, modern corpus planning, LPP and technology, and language revitalisation.

Keywords Language policy · Language planning · Epistemology · Theory

1 AN INTERDISCIPLINARY FIELD OF RESEARCH

Language policy and planning (LPP) is an interdisciplinary field, which relies on contributions from the humanities and social sciences, and thus demonstrates robust theoretical and epistemological diversity. While it emerged from sociolinguistics (Haugen 1964) and the sociology of language (Fishman 1972), LPP research is characterised by a marked

multidisciplinarity and interdisciplinarity with researchers leveraging the theories and methods from their disciplinary homes to examine LPP documents, processes, and impacts. Examples include (socio)linguistics (Hult 2010), economics (Grin and Vaillancourt 2015; Gazzola and Wickström 2016), political theory (May 2008; Sontag 2009), legal theory and analysis (Kochenov and de Varennes 2015), educational linguistics (Johnson 2013), and anthropology (Hornberger 1988). This multidisciplinarity was the motivation for the annual *Multidisciplinary Approaches in Language Policy and Planning Conference* founded in 2012 and convened by Thomas Ricento. As well, edited volumes by Hult and Johnson (2015), and Grin et al. (2022) represent diverse research methods and approaches in LPP, including chapters on political theory, legal theory, ethnography, corpus linguistics, economics, media studies, and demography (among others).

Early language planning research in the 1960s and 1970s (sometimes called 'classic LPP') was marked by objectivism and a growing yet inchoate batch of theoretical and methodological tools (e.g., Rubin and Jernudd 1971). At that time, the goal of LPP was mainly to help policy makers standardise and modernise indigenous languages in postcolonial settings and to select one or more official languages (typically the languages of the former colonial powers) to promote economic development and nation building (Fishman et al. 1968; Fishman 1974). Solutions for language problems were the focus—for example, what status to grant colonial languages in postcolonial nations—and language planning frameworks, models, and conceptualisations were developed and proposed. Language planning definitions focused on problem-solving and deliberate interventions. For example, Rubin and Jernudd (1971) argue that

> language planning is *deliberate* language change; that is, changes in the system of language code or speaking or both that are planned by organizations that are established for such purpose or given a mandate to fulfil such purpose. Language planning is focused on problem-solving and is characterised by the formulation and evaluation of alternatives for language problems to find the best (or optimal, most efficient) decision. (Rubin and Jernudd 1971, xvi, emphasis in the original)

It was an experimental time in the field, with scholars relying on the tools from various disciplines to analyse language planning. As Fishman (1977: 33) reflects, the primary aim of early language planning projects was

to demonstrate "the feasibility of studying language planning processes" rather than testing specific hypotheses or evaluating the effectiveness of particular research methods. This early work engendered edited volumes like *Language Problems of Developing Nations* (Fishman et al. 1968), *Can Language Be Planned?* (Rubin and Jernudd 1971), *Language Planning Processes* (Rubin et al. 1977), and the foundation of the first journal in the field in 1969, i.e., *La monda lingvo-problemo* ("The world language problem" in Esperanto), which in 1977 was renamed *Language Problems & Language Planning* (and still exists under this name). Most of the contributions were conceptual proposals, descriptive accounts of language planning projects, strategies for implementing language planning, and historical investigations of particular contexts and communities. The goals and methods of early language planning research were influenced by an interest in excluding sociopolitical variables in the objective science of language planning.

Responding to post-enlightenment objectivism, the focus on the rational individual, and the view of LPP as a technical and instrumental problem-solving task, Tollefson (1991) proposed the historical-structural approach, which incorporated critical social theory and introduced a different epistemological orientation to the field. Contrasted with earlier optimistic assumptions that language planning could solve language problems, Tollefson focused on how language planning led to systemic inequality. His critique of early language planning research, accompanied by an innovative new vision for LPP research, marked an epistemological and theoretical turning point. A number of scholars have examined power issues hidden behind language policies that accompanied attempts at modernisation and economic development of developing countries (e.g., Phillipson 1992). Multiple publications have taken up this critical perspective, which is also a defining feature of the current iteration of the journal *Language Policy*.

These contributions gave rise to Critical Language Policy (CLP), which is sometimes portrayed as neo-Marxist (Fishman 1994), since it relies on critical social theorists like Bourdieu (1977), Habermas (1973), and Giddens (1979), and/or postmodernists like Foucault (1972) (see Grin 2022 for an overview of the differences and commonalities between these different approaches). Neo-Marxism and postmodern critical social theory focus on how power circulates in society, often outside of conscious human control, thus minimising the role of human agency in social processes. For example, Bourdieu argues that the agent is "not the

producer and has no conscious mastery over their actions" (p. 79). Similarly, Foucault (1978: 95) argues that while counter-discourses are representative of the inevitability of resistance, "this resistance is never in a position of exteriority in relation to power". Critiques of critical language policy scholarship (e.g., Ricento and Hornberger 1996), emphasise a more robust consideration of human agency in LPP processes. Furthermore, reacting to the monolithic determinism and hegemony in language policy creation, Hornberger (2005) focuses on how language policies play an important role in providing implementational and ideological spaces for multilingualism (cf. Fishman 1994). An empirical focus on how individuals create, interpret, and appropriate language policies across diverse contexts and levels of institutional authority helped engender an epistemological orientation that both accepted the impact of the researcher on the research context and based claims on empirical research of LPP processes.

Johnson and Ricento (2013) describe the beginning of the 1990s as a third phase in the history of LPP research, which is characterised by an increasing reliance on discursive and ethnographic methodological approaches (Barakos and Unger 2016; Pennycook 2006; Tollefson and Pérez Milans 2018; Wodak 2005). The development and the application of ethnographic methods are notable during this phase, shifting the focus of language planning from the macro- to the meso- and micro-levels, e.g., within specific speech communities. The ethnography of language policy reveals "the agents, contexts, and processes across the multiple layers of language policy creation, interpretation, and appropriation" (Johnson 2013: 44). Contributions in this area often study the tensions and contradictions between the general objectives of a formal language policy and the observed practices on the ground, especially in education (see Hornberger and Johnson 2011; McCarty 2015).

While Johnson and Ricento (2013) portray the different eras as "phases" in the evolution of LPP research (see also Lo Bianco 2010), in reality many different approaches evolved in tandem, both together and independently, including classic approaches, which continue to this day. Starting in the 1990s and increasingly since the 2000s, for example, a distinct research tradition with significant connections to early LPP scholarship was developed. Researchers from economics, rational choice theory, and public policy studies reintroduced, updated, and expanded concepts and tools from policy analysis into LPP (see Vaillancourt 1985; Pool 1991; Grin 2003; Wickström 2007; Gazzola 2014). This tradition

questions the neo-Marxist argument that LPP is primarily or exclusively focused on the hegemonic promotion of the ruling class's interests, and argues that official government-driven LPP can be democratic, diversity-oriented, and aimed at reducing inequalities. This raises the question of what policy design theories and implementation instruments, as well as evaluation methods, are best suited to promote multilingualism and social justice, and how to assess the costs and the benefits of different LPP measures. For example, efforts to produce a national, comprehensive, multi-interest language policy to address language problems of minority and majority communities (as opposed to entrenching inequalities) through a combination of bottom-up and top-down processes were done in Australia in the 1980s (Lo Bianco 1987).

Another interesting development in the study of language policy during the last two decades is the result of political philosophers' and political scientists' work. Starting in the 1990s, but more intensively since the 2000s (e.g., Kymlicka and Patten 2003; Van Parijs 2011), political philosophers and theorists have focused on language in society and the corresponding policies. One focus within the debate has been the moral justifications underpinning language policy, especially in the area of the rights of linguistic minorities and equity in international communication. The debate among political philosophers has been structured along the liberal/multiculturalism divide—that is, following either the liberal tradition represented by philosophers such as John Rawls (1971), or the communitarian/multiculturalist tradition inspired by the work of thinkers such as Charles Taylor (1992). This research area is sometimes referred to as 'linguistic justice' (see Morales-Gálvez and Riera-Gil 2019; Alcalde 2018; De Schutter 2007 for extensive reviews; see also Branchadell 2005 for a discussion about the relationships between this strand of literature and research in LPP from a public policy perspective). The issue of power in language policy, in particular, is the object of an important tradition of papers published by political scientists interested in the politics of language (see Laponce 2006; Cardinal and Sonntag 2015; Kraus 2018; May 2011; Ives 2015). These contributions find their roots in research carried out by some major political scientists such as Karl Deutsch, Stein Rokkan, and Jean Laponce in the areas of multilingualism and linguistic diversity between the 1950s and 1980s (see Laponce 2004 for a survey).

As a result, LPP has become not only multidisciplinary but also increasingly interdisciplinary, or even transdisciplinary (Halliday 1990). The chapters in this book reflect this diversity, and they present some of the

distinct theoretical and epistemological positions within LPP research, as well as some recent theoretical developments, notably at the intersection between LPP and information and communication technology. The book deliberately focuses on the contributions of academic research, but this of course does not mean denying the relevance of the theoretical and epistemological input of practitioners in language policy (a point to which we will return in the conclusions).

Chapter 2 by Johnson considers the epistemological and theoretical foundations of critical empirical approaches in LPP research, with particular attention to ethnography and discourse analysis. It traces early language planning research from the 1960s and 1970s to developments in the 1980s and 1990s that reflected a broader trend in sociolinguistics and the language sciences more generally, towards critical conceptualisations that focused on power in language and language learning. A growing critical perspective was, then, influential for a new generation of LPP scholars in the 2000s, 2010s, and 2020s who leveraged disciplinary research methodologies to study LPP processes. Much of this work has focused on how language policies impact educational opportunity and equity for speakers of minoritised and marginalised languages. Chapter 2 considers the findings from ethnographic and discourse analytic studies and the value of these contributions to the field. Both the criticisms of critical approaches and the limitations of ethnography and discourse analysis are discussed.

Chapter 3 by Gazzola presents and discusses the public policy approach to LPP. The chapter critically examines the idea according to which everyone (including single individuals) can make language policy, and it emphasises the differences between (individual) language practices and (public) language policy. It argues that the government (at different levels, from the local to the national, and even supranational) is the central agent in the language policy process. In the chapter, therefore, language policy is presented as a particular form of public policy that can be examined using the policy cycle model, which is the standard analytical framework in public policy studies. The model comprises several stages, namely the emergence of a language issue in society, followed by agenda-setting, policy formulation and adoption, implementation, and finally evaluation. The chapter presents these five stages of the cycle, the relationships between them, and clarifies how the various disciplines involved in LPP can contribute to their study. It also presents some concepts that are central in the design of language policies, such as programme

theory, policy instruments, and indicators. Finally, the chapter discusses the limits of the model. The chapter emphasises the pragmatic aspects of LPP and the need to engage with decision-makers' practical concerns. In this sense, it contributes to restoring the theoretical and epistemological links between LPP research and public policy studies that have gradually weakened after the 'critical turn' in LPP.

Chapter 4 by Gobbo advocates for a reconsideration of corpus planning in LPP research as an empirical testbed of the effectiveness of language policy actions and the feasibility of scenario formulations. The main argument is that any language policy eventually influences the materiality of the languages involved, in particular minority languages, but also major languages in contact. Case studies should consider all the languages involved and what happens in their respective corpus, at least at the level of the lexicon, but also on the level of morphology and sometimes even phonology. The chapter illustrates the method of the levels of abstraction and the procedure to perform the epistemic analysis of LPP case studies, adapted respectively from the Philosophy of Information and Computer Ethics. Case studies of the (re)vitalisation of Esperanto and Hebrew are compared, which are offered as an example to illustrate the epistemological standpoints, so as to avoid biases in formulating language policies, as best as possible.

Chapter 5 by Leoni de León approaches the convergence of Natural Language Processing (NLP), Theoretical Linguistics (TL), and Language Policy and Planning (LPP). In order to take advantage of resources and opportunities, 30 fundamental questions are given on methodological and ethical matters which will help define the profile of any project and improve the chances of success. The attitudes, expectations, and needs of the language communities should be considered in any technological language project, in order to offer better solutions. In recent years, the intertwining between language and technology has increased, and the presence of Information Technology is overwhelming. Therefore, it is important to be conscious that the advantages that technology offers for dominant languages cannot necessarily be successfully reproduced for minority or endangered languages. In addition, we must overcome the current situation in which standards are de facto defined by private entities following their own interest. As a consequence, the interest of language communities, especially minority languages and endangered languages, are not necessarily well served. The development of language applications requires the empowerment of the technology by the linguistic communities themselves.

2   Language Policy and Language Planning

Debated and debatable definitions proliferate throughout this book and, here, a discussion of language policy and language planning is necessary. The terms 'language policy' and 'language planning' are sometimes used as synonyms, but they have unique origins and evolutions. Haugen introduced the term language planning in 1959, defining it as "the activity of preparing a normative orthography, grammar, and dictionary for the guidance of writers and speakers in a non-homogeneous speech community" (Haugen 1959: 8). What Haugen describes here would become known as corpus planning which includes activities related to the manipulation of the forms of a language. While many language planners and scholars focused on *corpus planning*, others studied how a society could best allocate functions and/or uses for particular languages, known as *status planning*, a distinction introduced by Kloss (1969). Cooper (1989) introduced the concept *acquisition planning*, which describes all attempts to capture language teaching and other educational activities designed to increase the users or uses of a language.

Most would agree that language policy and language planning are closely related but different activities. Some argue that language planning subsumes language policy (Kaplan and Baldauf 1997), while others argue that language policy subsumes language planning (Schiffman 1996). The term 'language planning' is the first one to appear; however, the term 'language policy'—along with others 'language management'—was introduced to complement 'language planning', by those who thought the definitions unsatisfactory (Phillipson 1992: 86). Some authors argue that 'language planning' addresses the internal changes of a language steered by conscious intervention, while 'language policy' is more concerned with the roles of language and its speakers in a given political, social, and cultural context (see Gobbo, this volume). Current research often combines the terms resulting in the acronym LPP. Experience in the field has revealed that both aspects—i.e., language policy and language planning—should be addressed in tandem in order for LPP efforts to be effective. Because of its heterogeneous origin, LPP has subsequently developed in non-linear ways along different theoretical and methodological approaches. Still, the order of the terms—language policy and planning or language planning and policy—probably indexes a preference by the author in question. We argue that future research in LPP should be explicit about the epistemological foundation for particular

research projects in order to maximise clarity, minimise misunderstandings, and encourage cross-fertilisation across disciplines, so as to enlarge its epistemological horizon.

3   Origins and Target Audience of the Book

In 2017 and 2021, at the initiative of Michele Gazzola and Antonio Leoni de León, the Institute of Linguistic Research of the University of Costa Rica organised and funded a Summer School on *Linguistic Policy, Language Planning and Evaluation*. The event benefited also from the collaboration and the financial support of Esperantic Studies Foundation (http://esperantic.org). The authors gratefully thank both organisations for their support.

The main purpose of the two editions of the summer school was to consolidate a space for the discussion and dissemination of current issues of language policy in Latin America, as well as to promote international scientific collaborations. There was a need to bring together researchers, teachers, students, and professionals from public institutions and non-governmental organisations in the region. The historically opposing forces over several centuries and the different forms of intercommunity contact in Latin America have produced the current linguistic panorama, which combines the imposition of dominant languages, the resistance to persistent attempts at repression, and linguistic and cultural miscegenation. In Hispanic linguistics, indeed, LPP has occupied an important place since the nineteenth century, in particular in relation to the efforts of the newly independent Latin American nations to promote their local Spanish varieties as a means of marking each country's singularities against the former colonial power, but at the same time claiming a pan-Hispanic identity.[1]

---

[1] The case of Mexico is particularly interesting (Cifuentes and Ros 1993). The linguistic policies of the nineteenth century were part of a much broader process in the building of national identities (Marcilhacy 2018). A fundamental figure in this sense is Andrés Bello, who through the publication of several grammatical works affirmed this spirit of independence and Hispanic union. It is not surprising, then, that language planning and language policy have been a recurring theme among Hispanic linguists and philologists, who have proposed, at various times, national and international standards for the Spanish language that have led to the recent pan-Hispanic trend of recognizing the pluricentrism of the Spanish language (Amorós-Negre and Prieto de los Mozos 2017; Blas Arroyo 1998–1999; Moreno Fernández 1992).

The first edition of the summer school in 2017 covered general topics such as language policy evaluation, language planning in education, indigenous languages and language shift, sign languages, and constructed languages as an extreme form of language planning. At the conclusion of the event, given the depth and diversity of the topics, the invited instructors as well as the organisers, i.e., David Cassels Johnson, Federico Gobbo, Michele Gazzola, and Antonio Leoni de León, felt the need to share their views through a monograph that accounts for the diversity of perspectives and the complexity and dynamism of the LPP field, and to provide a short and accessible book deliberately designed for graduate and postgraduate students. This is how the present volume began to take shape, and we are delighted that the time has come for its publication.

References

Alcalde, Javier (2018). Linguistic justice: An interdisciplinary overview of the literature. In Michele Gazzola, Torsten Templin, and Bengt-Arne Wickström (eds.), *Language Policy and Linguistic Justice: Economic, Philosophical and Sociolinguistic Approaches*, pp. 65–149. Berlin: Springer.

Amorós-Negre, Carla, and Prieto de los Mozos, Emilio (2017). El grado de pluricentrismo de la lengua Española. *Language Problems & Language Planning*, 41 (3), pp. 245–264. https://doi.org/10.1075/lplp.00004.amo

Barakos, Elisabeth, and Unger, Johann (eds.) (2016). *Discursive Approaches to Language Policy*. Basingstoke: Palgrave Macmillan.

Blas Arroyo, José Luis (1998–1999). Desarrollos de la plani-ficación lingüística en el mundo hispánico con especial atención a los contextos español y lati-noamericano. *Revista española de lingüística aplicada*, ISSN 0213–2028, 13, pp. 315–344.

Bourdieu, Pierre (1977). *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.

Branchadell, Albert (2005). *La moralitat de la política lingüística: Un estudi comparat de la legitimitat liberaldemocràtica de les polítiques lingüístiques del Quebec i Catalunya*. Barcelona: Institut d'Estudis Catalans.

Cardinal, Linda, and Sonntag, Selma K. (eds.) (2015). *State Traditions and Language Regimes*. Montreal: McGill-Queen's University Press.

Cifuentes, G. Bárbara, and Ros, María del Consuelo R. (1993). Oficialidad y planificación del español: dos aspectos de la política del lenguaje en México durante el siglo XIX. *Iztapalapa Revista de Ciencias Sociales y Humanidades*, 13 (29), pp. 135–146. https://revistaiztapalapa.izt.uam.mx/index.php/izt/article/view/1162

Cooper, Robert L. (1989). *Language Planning and Social Change*. Cambridge: Cambridge University Press.

De Schutter, Helder (2007). Language policy and political philosophy: On the emerging linguistic justice debate. *Language Problems & Language Planning*, 31 (1), pp. 1–23.

Fishman, Joshua A. (1972). The sociology of language: An interdisciplinary social science approach to language in society. In Joshua A. Fishman (ed.), *Basic Concepts, Theories and Problems: Alternative Approaches*, pp. 217–404. Berlin and Boston: De Gruyter Mouton. https://doi.org/10.1515/978311141750 9-005

Fishman, Joshua A. (1974). *Advances in Language Planning*. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783111583600

Fishman, Joshua A. (1977). Selected dimensions of language planning: A comparative analysis. In Joan Rubin, Björn H. Jernudd, Jyotirindra DasGupta, Joshua A. Fishman, and Charles A. Ferguson (eds.), *Language Planning Processes*, pp. 195–214. Berlin and Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110806199.195

Fishman, Joshua A. (1994). Critiques of language planning: A minority languages perspective. *Journal of Multilingual and Multicultural Development*, 15 (2–3), pp. 91–99.

Fishman, Joshua A., Ferguson, Charles A., and Das Gupta, Jyotirindra (eds.) (1968). *Language Problems of Developing Nations*. New York and London: Wiley.

Foucault, M. (1972). *The Archaeology of Knowledge and the Discourse on Language*. New York: Pantheon Books.

Foucault, M. (1978). *The History of Sexuality: An Introduction*. New York: Random House.

Gazzola, Michele (2014). *The Evaluation of Language Regimes: Theory and Application to Multilingual Patent Organisations*. Amsterdam: John Benjamins.

Gazzola, Michele, and Wickström, Bengt-Arne (eds.) (2016). *The Economics of Language Policy*. Cambridge, MA: MIT Press.

Giddens, Anthony. (1979). *Central Problems in Social Theory: Action, Structure and Contradiction in Social Analysis*. Berkeley: University of California Press. https://doi.org/10.1007/978-1-349-16161-4

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

Grin, François (2022). Progress in LPP: Towards an assessment of challenges from critical perspectives. *Sociolinguistica*, 36 (1–2), pp. 85–97. https://doi.org/10.1515/soci-2022-0002

Grin, François, Marácz, László, and Pokorn, Nike K. (eds.) (2022). *Advances in Interdisciplinary Language Policy*. Amsterdam: John Benjamins.

Grin, François, and Vaillancourt, François (2015). The economics of language policy: An introduction to evaluation work. In Francis M. Hult, and David Cassels Johnson (eds.), *Research Methods in Language Policy and Planning: A Practical Guide*, pp. 118–129. Hoboken: Wiley-Blackwell.

Habermas, Jürgen (1973). *Theory and Practice*. Boston, MA: Beacon Press.

Halliday, M. A. K. (1990). New ways of meaning: The challenge to applied linguistics. *Journal of Applied Linguistics*, 6, pp. 7–36.

Haugen, Einar. (1959). Planning for a standard language in Norway. *Anthropological Linguistics*, 1 (3), 8–21.

Haugen, Einar (1964). Linguistics and language planning. In William Bright (ed.), *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference*. Berlin and Boston: De Gruyter Mouton, 2015, pp. 50–71. https://doi.org/10.1515/9783110856507-006

Hornberger, N. H. (1988). *Bilingual Education and Language Maintenance*. Dordrecht, Holland: Foris Publications.

Hornberger, Nancy H. (2005). Frameworks and models in language policy and planning research by frameworks and models in language policy and planning. In Thomas Ricento (ed.), *An Introduction to Language Policy: Theory and Method*, pp. 24–41. Oxford: Blackwell.

Hornberger, N. H, and Johnson, D. C. (2011). The ethnography of language policy. In T. L. McCarty (ed.), *Ethnography and Language Policy*, pp. 273–289. New York: Routledge.

Hult, Francis M. (2010). Analysis of the language policy discourses across the scales of space and time. *International Journal of the Sociology of Language*, 202, pp. 7–24.

Hult, Francis M., and Johnson, David Cassels (eds.) (2015). *Research Methods in Language Policy and Planning: A Practical Guide*. Hoboken: Wiley-Blackwell.

Ives, Peter (2015). Global English and the limits of liberalism: Confronting global capitalism and challenges to the nation state. In Thomas Ricento (ed.), *Language Policy and Political Economy: English in a Global Context*, pp. 48–71. Oxford and New York: Oxford University Press.

Johnson, David Cassels (2013). *Language Policy*. Basingstoke: Palgrave Macmillan.

Johnson, D. C., and Ricento, T. (2013). Conceptual and theoretical perspectives in language planning and policy: Situating the ethnography of language policy. *International Journal of the Sociology of Language*, 219, pp. 7–21. https://doi.org/10.1515/ijsl-2013-0002

Kaplan, Robert B., and Baldauf, Richard B. (1997). *Language Planning: From Practice to Theory*. Clevedon: Multilingual Matters.

Kloss, H. (1969). *Research Possibilities on Group Bilingualism: A Report*. Quebec: International Center for Research on Bilingualism.

Kochenov, D., and de Varennes, F. (2015). Language and law. In F. Hult and D. Johnson (eds.), *Research Methods in Language Policy and Planning*. Malden, MA: Wiley.

Kraus, Peter A. (2018). From glossophagic hegemony to multilingual pluralism? Re-assessing the politics of linguistic identity in Europe. In Peter A. Kraus, and François Grin (eds.), *The Politics of Multilingualism. Europeanisation, Globalisation and Linguistic Governance*, pp. 89–110. Amsterdam: John Benjamins.

Kymlicka, Will, and Patten, Alan (eds.) (2003). *Language Rights and Political Theory*. Oxford: Oxford University Press.

Laponce, Jean (2004). Language and politics. In Mary Hawkesworth and Maurice Kogan (eds.), *Encyclopedia of Government and Politics*, pp. 587–602. London: Routledge.

Laponce, Jean (2006). *Loi de Babel et autres régularités des rapports entre langue et politique*. Quebec: Presses de l'Université Laval.

Lo Bianco, Joseph (1987). *National Policy on Languages*. Canberra: Australian Government Publishing Service.

Lo Bianco, Joseph (2010). Language policy and planning. In Nancy H. Hornberger and Sandra Lee McKay (eds.), *Sociolinguistics and Language Education*, pp. 146–176. Clevedon: Multilingual Matters.

Marcilhacy, David (2018). *Une histoire culturelle de l'hispano-américanisme (1910–1930): L'Espagne à la reconquête d'un continent perdu*. Thèse de doctorat. Sciences de l'Homme et Société. Sorbonne Nouvelle—Paris 3. https://hal.archives-ouvertes.fr/tel-01879267v1

May, Stephen (2011). *Language and Minority Rights: Ethnicity, Nationalism and the Politics of Language* (2nd ed.). London: Routledge.

McCarty, Teresa L. (2015). Ethnography in language planning and policy research. In Francis M. Hult and David Cassels Johnson (eds.), *Research Methods in Language Policy and Planning: A Practical Guide*, pp. 81–93. Hoboken: Wiley-Blackwell.

Morales-Gálvez, Sergi, and Riera-Gil, Elvira (2019). Què són polítiques lingüístiques justes? Els paradigmes actuals de la justícia lingüística. *Revista d'Estudis Autonòmics i Federals*, 30, pp. 25–56.

Moreno Fernández, F. (1992). Norma y prestigio en el español de América. Apuntes para una planificación de la lengua española. *Revista De Filología Española*, 72 (3/4), pp. 345–360. https://doi.org/10.3989/rfe.1992.v72.i3/4.565

Pennycook, Alastair (2006). *Global Englishes and Transcultural Flows*. London: Routledge.

Phillipson, Robert (1992). *Linguistic Imperialism*. Oxford: Oxford University Press.

Pool, Jonathan (1991). The official language problem. *American Political Science Review*, 25 (2), pp. 485–514.

Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Ricento, T., and Hornberger, N. H. (1996). Unpeeling the onion: Language planning and policy and the ELT professional. *TESOL Quarterly*, 30 (3), pp. 401–427.

Rubin, Joan, and Jernudd, Björn H. (1971). Introduction: Language planning as an element in modernization. In Joan Rubin and Björn H. Jernudd (eds.), *Can Language be Planned? Sociolinguistic Theory and Practice for Developing Nations*, pp. xiii–xxiv. Honolulu: The University Press of Hawaii.

Rubin, Joan, Jernudd, Björn H., DasGupta, Jyotirindra, Fishman, Joshua A., and Ferguson, Charles A. (eds.). (1977). *Language Planning Processes*. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110806199

Schiffman, Harold F. (1996) *Linguistic Culture and Language Policy*. London: Routledge.

Sontag, S. (2009). Linguistic globalization and the call center industry: Imperialism, hegemony, or cosmopolitanism? *Language Policy*, 8, pp. 5–25.

Taylor, Charles (1992). *Multiculturalism and the Politics of Recognition*. Princeton: Princeton University Press.

Tollefson, James W. (1991). *Planning Language: Planning Inequality*. New York: Longman.

Tollefson, J. (2011). Language planning and language policy. In R. Mesthrie (ed.), *The Cambridge Handbook of Sociolinguistics*, pp. 357–376. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511997068.027

Tollefson, James William, and Pérez Milans, Miguel (2018). Research and practice in language policy and planning. In James William Tollefson and Miguel Pérez Milans (eds.), *The Oxford Handbook of Language Policy and Planning*. Oxford: Oxford University Press.

Vaillancourt, François (ed.) (1985). *Économie et langues*. Québec: Conseil de la langue française.

Van Parijs, Philippe (2011). *Linguistic Justice for Europe and for the World*. Oxford: Oxford University Press.

Wickström, Bengt-Arne (2007). Fairness, rights, and language rights: On the fair treatment of linguistic minorities. In Pio Baake and Rainald Borck (eds.), *Public Economics and Public Choice: Contributions in Honor of Charles B. Blankart*, pp. 81–102. Berlin: Springer.

Wodak, Ruth (2005). Linguistic analyses in language policies. In Thomas Ricento (ed.), *An Introduction to Language Policy: Theory and Methods*, pp. 170–193. Oxford: Blackwell.

Critical Empirical Approaches in Language Policy and Planning

David Cassels Johnson

Abstract This chapter considers the epistemological and theoretical foundations of critical empirical approaches in language policy and planning (LPP) research. Developments in the 1980s and 1990s reflected a broader trend in sociolinguistics, and the language sciences more generally, towards critical conceptualisations that focused on power in language, language learning, and language policy processes. This critical perspective was influential for a new generation of LPP scholars in the 2000s and 2010s. Much of this work has focused on how language policies impact educational opportunity and equity for speakers of minority languages. After the historical overview of the epistemological and theoretical foundations, this chapter considers findings from ethnographic and discourse analytic studies and the value of these contributions to the field. Finally, the chapter considers the future role of critical empirical approaches in the field of LPP. In particular, debates about agency and structure, the macro-micro dialectic as a theoretical took, and researcher positionality are considered, in the light of theoretical developments and empirical findings.

1   INTRODUCTION

This chapter considers the epistemological and theoretical foundations of
critical empirical approaches in LPP research, with particular attention to
ethnographic and discourse analytic traditions. Ricento (2000) organises
the historical development of language policy and planning (LPP) into
three phases: (1) the origins of the field and classic language planning
research; (2) expanding frameworks in the 1980s, with the introduc-
tion of new disciplinary theories and methodologies; and (3) Critical
Language Policy. In a re-examination of the field, Johnson and Ricento
(2013) propose ethnography of language policy as central to a fourth
phase of LPP research. This chapter follows that historical analysis but
further explores the epistemological, theoretical, and empirical contribu-
tions and debates during this so-called fourth phase. Both the criticisms
and limitations of critical approaches are discussed, and particularly influ-
ential contributions are highlighted. Finally, the chapter considers the
future role of critical empirical approaches in the field and, in partic-
ular, its role in debates about agency and structure, the macro-micro
dialectic, and researcher positionality. The chapter ends with an argu-
ment about the importance of a transdisciplinary approach going forward,
which complements the foci discussed in the other chapters.

2   EPISTEMOLOGICAL AND THEORETICAL
FOUNDATIONS IN LANGUAGE PLANNING

All research is grounded in an epistemological paradigm, even if not
discussed or recognised explicitly in the research report. Epistemology
is the branch of philosophy focused on the nature of knowledge and
provides the philosophical foundation for the theoretical framework,
methodology, and methods in a research project. Epistemological ques-
tions include: how do we come to know what we know? How do we ask
and answer research questions? How can we be sure that our answers to
those questions are accurate, valid, or valuable?

The connections between the historical development of epistemolog-
ical positions and modern research methods are sometimes lost and,

here, I consider the term "empirical" as it connects to historical developments and modern data collection. Empiricism, as a philosophy of science promoted by philosophers like John Locke, describes the position that knowledge is developed through observation and experience. Innate ideas, religious and cultural traditions, metaphysical beliefs, and a priori assumptions are, therefore, not considered good evidence for answering research questions. Growing out of the more general intellectual movement, Logical Empiricism was a philosophical movement that arose much later, in the early twentieth century, and while it was not defined by a unitary doctrine, there was a common concern with scientific methodology. It was preceded by, and closely connected to, Logical Positivism, an intellectual movement connected to scholars like Wittgenstein, which grew out of the general position of positivism, often associated with August Comte. Among all these movements, there is a focus on using scientific methods for the study of humanity, and in particular, applying research methods from the natural sciences to sociology. Comte, for example, rejected supernatural explanations, elevated empirical observation, and proposed a hierarchy of sciences, with mathematics at the top. As a research philosophy, positivism promotes universality in research methods and strict scientific guidelines, attempts to systematically control the variables in a study, and rejects metaphysical logic. However, as was the case during the time of its development, the term positivism is "better known through the enemies of that mode of thinking than through its friends" (Mill 1866) and, today, is often simply a term of abuse used by anti-positivist scholars. Furthermore, as Crotty (1998) argues, it is a mistake to equate Comte with a mathematical approach to research, controlled experimentation, or even positivism. While Comte used the term "positive science", he did not envision himself as a positivist. Crotty sees the origins of positivism instead, in Francis Bacon.

The emphasis on empirical data collection continues to dominate many research traditions across the academy, yet conceptions of "empirical" vary. For example, while there is sometimes a tendency to equate empirical with data collection techniques found in the natural sciences or, even less accurately, as a synonym for "quantitative", this conflicts with both historical and modern definitions. Certainly, experimental data are empirical, but so are observations and interviews, for example. The term itself comes from the Greek word "experience", reflecting its relationship to direct experience, as opposed to unobservable or theoretical justification. Nevertheless, while researcher intuition is not empirical evidence, there is

wide recognition in the natural and social sciences that it plays a crucial role in the research process. Finally, while empirical data collection is often conflated with objectivity in data collection—and thus would exclude critical research traditions, for example—this is a debatable position.

Divisions in, and descriptions of, epistemological orientations vary widely. For example, some argue that research traditions like positivism or postmodernism are epistemologies (Madison 2012), others consider them theoretical orientations (Crotty 1998), while others describe them as paradigms (Glesne 2016). Crotty (1998) proffers a helpful and influential framework, which distinguishes between only three major epistemological orientations that guide academic research—objectivism, constructionism, and subjectivism—out of which theoretical orientations like positivism and postmodernism fall. Objectivism is the belief that objects exist in the world as meaningful entities, independent of human consciousness; that is, they have meaning with or without a human observer, and it is the researcher's mission to discover the inherent meaning. Subjectivism, on the other hand, argues the oppositive—objects have no meaning without a human observer, who imposes meaning onto the entity under consideration. Lying in the middle, so to speak, is constructionism, which contends that meaning is developed out of the interaction between the human observer and that which is observed. Note that none of these positions refer to the *existence* of objects, because the nature of being (not knowledge) is an ontological question. Therefore, a rock on some distant planet may exist without a human observer (which is a realist ontological position), but the subjectivist might still argue that the rock has no *meaning* without a human observer—these positions are not contradictory. Of course, while referring to inanimate objects helps illustrate the philosophical positions, the questions become more difficult when we consider human beings as the "objects" under consideration.

Some theoretical orientations and methodological traditions naturally align with particular epistemologies. For example, positivism neatly aligns with objectivism. Yet, the evolution to post-positivism and the acceptance that the practice of scientific research is often influenced by the subjectivity of the researcher perhaps muddies the waters. For example, the ontologically bewildering double-slit experiment in quantum mechanics has led some to argue that quantum theory "should be interpreted in an observer-dependent way" (Proietti et al. 2019). Still, within post-positivist social science research, experimental studies with random assignment and generalisability are the gold standard. On the other hand, critical traditions like postmodernism reject modernist notions of what

counts as "science", and therefore, it is difficult to imagine a postmodern research project guided by objectivism. However, other research traditions have cut across the epistemological spectrum. For example, there is a strong history of objectivist ethnographic research, influenced by naturalism in the biological sciences, in which the human observer attempts to unobtrusively observe the behaviour of others as a "fly on the wall" (Roman 1993). Even though the emphases in ethnographic projects have shifted over the years—with contemporary critical ethnographic approaches aligning with constructionism or subjectivism (e.g., Madison 2012)—dilemmas in participant observation still remain a bedrock data collection challenge within ethnographic studies.

As mentioned in the introduction, early language planning research in the 1960s and 1970s was grounded in objectivism and utilized a growing yet inchoate batch of methodological tools (e.g., Rubin and Jernudd 1971). Solutions for language problems were the focus—for example, what status to grant colonial languages in postcolonial nations— and language planning frameworks, models, and conceptualisations were developed. It was an experimental time in the field, with scholars leveraging discipline-specific tools to analyse language planning. This early work engendered edited volumes like *Language Problems of Developing Nations* (Fishman et al. 1968), *Can Language Be Planned?* (Rubin and Jernudd 1971), and *Language Planning Processes* (Rubin et al. 1977). Most of the contributions were conceptual proposals, descriptive accounts of language planning projects, theoretical proposals for language planning processes, and historical investigations of particular contexts and communities. The goals and methods of much early language planning research were influenced by an interest in excluding sociopolitical variables from the objective science of language planning.

Along with the development of the journal *La monda lingvo-problemo* in 1969, which in 1977 was renamed *Language Problems & Language Planning,* an important development in the field was the introduction of the *International Journal of the Sociology of Language* (IJSL) in 1974, which featured LPP research. In the first issue, Joshua Fishman writes that the sociology of language, as a field, must apply linguistic theories to real-world issues and he stresses the ethical responsibility of researchers: "We…all have a responsibility — to the weak, to the poor, to the uneducated, to the disadvantaged, to the discriminated, to the unpopular — to combine understanding, interpretation and application". While early language planning research may have been influenced by objectivism,

scholars like Fishman were still committed to an emancipatory research programme guided by tenants of social justice. The first article in the issue, also by Fishman, is an examination of language planning in Israel and, since then, the journal has been a crucial outlet for LPP research.

3   EXPANDING FRAMEWORKS AND EMPHASES

Developments in LPP research in the 1980s reflected a broader trend in sociolinguistics and the language sciences more generally, towards critical conceptualisations that foregrounded social aspects of language acquisition, practices, and learning. Outside of linguistics, critical social theorists like Michel Foucault and Pierre Bourdieu developed influential proposals about the role of language and power in social structures, which impacted scholarship across disciplines. While these influences would most notably take form in Tollefson's historical-structural approach (Tollefson 1991), their inchoate impact began to be seen in earlier work. As the field developed, the objectivism that characterised early LPP scholarship was called into question, a movement that was aligned with critical linguistics (Fowler et al. 1979), systemic functional linguistics (Halliday 1978), and developments in sociolinguistics (Hymes 1964). For example, Cobarrubias (1983) argues that corpus and acquisition planning are not "philosophically neutral" (p. 41) because status planning is inextricably linked to the functions of language in social institutions such as education and religion, and, furthermore, the actions of language planners themselves are tied to ideological principles with social consequences for the populations they impact.

Another essential contribution during the 1970s and 1980s comes from Arnold Leibowitz, a lawyer who examines English literacy and language policies and their impact on immigration, naturalisation, voting, and education in the U.S. (e.g., Leibowitz 1984). Foreshadowing the Critical Language Policy (CLP) approaches that would follow, Leibowitz reveals how language policy is a tool for discrimination, which provides a "palatable disguise for racist action" (p. 59). For example, the 1952 Naturalization Act added a literacy requirement to the already established speaking requirement for U.S. citizenship, which, Leibowitz argues, codified the courts' decisions that knowledge of English was necessary to prove assimilation and attachment to the constitution. Leibowitz further documents how the Know-Nothing Party was influential in convincing U.S. states to adopt a requirement for English literacy as a prerequisite for

suffrage, which was used to prevent Black Americans from voting. This was recognised by the U.S. Supreme Court in the 1960 decision of the United States vs. Mississippi (1960), in which it was argued that literacy requirements were specifically designed to affect voting rights. Wiley (1999) notes that the work of Leibowitz presaged the CLP approaches that would follow.

Other scholars considered language planning and policy as ideological processes. For example, Ruiz (1984) introduces an influential proposal, in which he lays out a tripartite set of orientations towards minority languages, which can take a language-as-problem, language-as-right, or language-as-resource orientation. Ruiz (1984: 2) characterises these 'orientations' as "basic to language planning in that they delimit the ways we talk about language and language issues... they help to delimit the range of acceptable attitudes toward language, and to make certain attitudes legitimate. In short orientations determine what is thinkable about language in society". The idea that language policies can hegemonically normalise ways of thinking, being, and/or educating, while concomitantly delimiting others, is aligned with critical social arguments from Foucault (1978) and echoed within critical LPP scholarship. Ruiz' arguments were indicative of a trend in the field, which was influenced by a movement away from objectivist research, that transcended individual disciplinary divisions and foregrounded power and inequality. Another example is the influential book *Language Planning and Social Change*, in which Cooper (1989) considers the sociopolitical impact of language planning, emphasising the possibility of social change: "Language planning, concerned with the management of change, is itself an instance of social change" (Cooper 1989: 164). Both Ruiz and Cooper foreshadowed the critical turn in LPP research, which would, in turn, influence a new generation of LPP scholars and set a decidedly different epistemological and theoretical course for the field.

There was a growing emphasis in the 1980s on how language policy could be used for discrimination and a final note about Neville Alexander is appropriate. Alexander helped found the National Liberation Front in South Africa and spent ten years in Robben Island prison with Nelson Mandela. After his release in 1974, he focused on multilingual education and linguistic human rights, founding the Project for the Study of Alternative Education in South Africa (PRAESA), which promotes multilingual education and language planning initiatives that promote the status of indigenous South African languages like isiXhosa. PRAESA

has successfully combined activism with research, contributing to the National Language Policy Framework and the South African Languages Bill (1999–2004). In his own scholarship, Alexander (1989: 62–63) argues that, while South African language policy is a tool of discrimination, the alternative is language planning "from below", which includes participation and consent by those affected by policy (students and parents) as well as a historical knowledge of past language planning and policy practices (i.e., what worked and what did not). I leave the final word with him: "Language policy can become an instrument to unify our people instead of being the instrument of division which, for the most part, it is today. We need to make a democratically conceived language policy an integral part of our programme for national unity and national liberation" (Alexander 1989: Preface).

4   Critical Language Policy

Early language planning research is criticised for being technocratic (Wiley 1999: 18), positivist (Ricento 2000: 208), or naïve (Kaplan et al. 2000), and critiques focus on the assumption that objectivist language planning, purported to be divorced form sociopolitical forces, could be effectively deployed to solve language problems. The criticism of early language planning research—that it ignored the ideological and political nature of language planning; that it focused on individual decisions instead of social, political, and economic forces; that it was overtly and overly positivist or technocratic—reflected a more general intellectual movement towards approaches influenced by postmodernism, post-structuralism, and critical social theory.

Tollefson was at the forefront of what Ricento (2000) describes as a third wave of LPP research, which is characterised by increasing attention to how language ideologies and discourses interact with LPP processes. Tollefson (1991) differentiates between what he calls the neo-classical approach to language planning and the historical-structural approach, which focuses on how historical and structural forces create and sustain systems of inequality. Tollefson (2016) portrays the historical-structural approach as part of a larger field of critical language studies, which foregrounds how relationships between language and power are created by sociocultural and sociolinguistic discourses (Fowler et al. 1979). Borrowing from Habermas (1973) and Foucault (1978), Tollefson (1991: 32, 35) proposes a new conceptualisation of language policy, describing it as "one mechanism by which the interests of dominant

socio-political groups are maintained and the seeds of transformation developed…The historical-structural model presumes that plans that are successfully implemented will serve dominant class interests". His critique of early language planning research, accompanied by an innovative new vision for language planning, was a watershed moment in the field and marked an epistemological and theoretical turning point, which would influence a new generation of LPP scholars.

In subsequent publications, Tollefson (2002, 2006, 2012) outlines the "critical" in CLP research, which (1) is critical of neo-classical language planning research; (2) is influenced by critical theory; (3) emphasises the relationships among language, power, and inequality, which are held to be central concepts for understanding language and society; and (4) entails social activism. While many researchers have taken up the first three aspects of CLP, the fourth has received less attention. This is not because LPP researchers are not committed to social justice, but because research and activism are often separated in ways that reify objectivist epistemologies. Countering this, Tollefson argues that language policy researchers are not only responsible for illuminating social hierarchies and processes of power, but also for changing those structures.

This demands a critical understanding of researcher subjectivity and positions the commitment to social justice as concomitant to, and reliant upon, a critical examination of interactions with participants: "[Critical language policy] researchers seek to develop a 'critical method' that includes a self-reflective examination of their relationship with the 'Others' who are the focus of research" (Tollefson 2002: 4).

While it was, and remains, influential (or perhaps because of its influence), the historical-structural approach has been criticised for being overly deterministic and underestimating the role of human agency (Ricento and Hornberger 1996), for not capturing language planning processes (Davis 1999), and for diminishing the potentially liberatory impact of language planning initiatives (Fishman 1994). Some of these criticisms are somewhat analogous to those levelled at critical social theory more generally, in which human agency appears to dissolve into dominant discourses that are all powerful and all consuming (Archer 2000). For example, while Foucault (1982) allows for "counter-discourses" in later writings, a consideration of human agency throughout his work is largely absent: any "resistance is never in a position of exteriority in relation to power" (Foucault 1978: 95). Discourse is both an effect and instrument of power, which has the ability to control human behaviour and thinking:

Discourses are "practices that systematically form the object of which they speak" (Foucault 1978: 49).

Within the critique of determinism proffered by Ricento and Hornberger (1996), they proposed the influential LPP onion metaphor to evoke the multiple layers through which language policy develops and they argue that LPP research has unsuccessfully accounted for activity in all layers. They emphasise the language policy power of the teacher, positioned at the centre of the onion, who exercises their power through pedagogical decisions—for example at one moment teachers may choose to incorporate multilingualism, positioning it as a resource, or they may choose not to, thereby closing potential spaces. Teachers are, therefore, not just policy implementers but policy *makers* (cf. Menken and García 2010). "We suggest that LPP is a multi-layered construct, wherein essential LPP components – agents, levels, and processes of LPP – permeate and interact with each other in multiple and complex ways as they enact various types, approaches, and goals of LPP" (Ricento and Hornberger 1996: 419–420). Expanding upon this a decade later, Hornberger and Johnson (2007) argue that the choices of educators are constrained by language policies, which tend to set boundaries on what is allowed and/or what is considered "normal", but the line of power does not flow linearly from the pen of the policy's signer to the choices of the teacher. The negotiation at each institutional level creates the opportunity for reinterpretations and policy manipulation. Local educators are not helplessly caught in the ebb and flow of shifting ideologies in language policies—they help develop, maintain, and change that flow.

In a separate line of critique, Fishman (1994: 96–97) warns against a myopic elevation of critical and ethnographic approaches to LPP research. The sanctification of ethnography and the "corresponding devilisation of other methods, smacks of Stalinism" and LPP research methods should be chosen based on "technically substantive rather than on trendy salvational grounds" (p. 97) (ostensibly portraying both ethnographic and critical approaches as "trendy").

Fishman (1994: 97) argues that language planning students, practitioners, researchers, and theoreticians are responsible for creating a better sociocultural reality for all those whose lives are impacted by language planning. Furthermore, Fishman (1994: 98) argues that "neo-Marxist critics…of language planning…never seem to go beyond their critique as decisively or as productively as they state their critique".

In an introduction to his edited volume, in which he responds to some of the criticism of CLP, Tollefson (2012) characterises two competing research approaches as the historical-structural and the creative public-sphere paradigms and argues that the distinction lies in the focus of the research, as opposed to some essential theoretical division. Indeed, much of the research epitomising the "empirical turn" relies on similar theoretical foundations as CLP research (and often the theoretical foundation *is* CLP). So, while the two approaches do offer competing conceptualisations of language policy texts, discourses, and practices, they share many theoretical and epistemological orientations.

5   Ethnography and Discourse Analysis

An emerging critical perspective was influential for a new generation of LPP scholars in the 2000s, 2010s, and 2020s. Much of this work has focused on schools, emphasising how language policies impact educational equity for speakers of minoritised languages. At the beginning of the century, two journals were introduced to the field, which helped cement these growing interests in LPP research. For the introduction to the first issue of *Current Issues in Language Planning*, Kaplan et al. (2000) argue that the naivety of early LPP research was plagued by Eurocentrism and a positivistic faith that "science" would solve the social and economic problems of former European colonies. They propose a systematic analysis that would consider "real language situations" in unrepresented or under-represented polities and promote an ecological perspective, borrowing from Haugen's (1972) ecology of language, which continues to guide many publications in CILP. Two years later, *Language Policy* was introduced, in which the editors, Spolsky and Shohamy (2002), emphasised "stronger theories" and "rigorous empirical studies" of cases, foreshadowing the definition that would be published in Spolsky's book (2004), which emphasises practices, ideology, and management as three aspects of language policy within a speech community.

Much of the empirical turn in LPP research has been propelled by ethnographic studies of language education and classroom discourse. All three LPP journals—*Language Problems and Language Planning*, *Current Issues in Language Planning*, and *Language Policy*—publish ethnographic and discourse analytic research although such manuscripts seem to be a more natural fit for *Language Policy* and, perhaps, *CILP*.

The focus of *LPLP* remains "large-scale language issues", while *CILP* focuses on LPP within polities, with special issues devoted to specific topics within the field. Individual articles "related to any area of language policy and planning" would include ethnographic and discourse analytic studies. A distinctive characteristic of *Language Policy* is its focus on both LPP research and educational policy and/or acquisition planning.

Inspired by both CLP approaches and Hymes' ethnography of communication, ethnographic research became an important tool for LPP research. Hornberger (1988), Davis (1994), King (2001), and McCarty (2002), for example, utilised ethnography to study LPP processes as they related to language maintenance and education for minoritised and indigenous language speakers. This body of research influenced and informed the ethnography of language policy, a research methodology that focuses on creation, interpretation, and appropriation of language policy texts and discourses across multiple levels and layers of LPP activity (Hornberger and Johnson 2007; Johnson 2009). Ethnographies of language policy are non-traditional in at least two important ways: first, the object of study is not a culture or a people, as would traditionally be the case, but a policy or policies. Second, traditional ethnographies are built on long-term participant observation among a particular community or within a particular context, but ethnographies of language policy often require data collection across diverse contexts and speech communities. Nevertheless, multi-sited ethnographic research is useful for studying policy (Levinson and Sutton 2001) and, even though the focus is "policy", human agents who are responsible for its creation, interpretation, and appropriation are always at the heart of ethnographic research. Ethnography has proven useful for uncovering LPP processes across very diverse contexts, a sampling of which includes language ideology and policy in Luxembourg (Davis 1999), bilingual education in Washington D.C. (Freeman 1998), indigenous language revitalisation in New Zealand (Hill and May 2011), linguistic landscapes and language ecologies in Sweden (Hult 2007), multilingual educational policy in Mozambique (Chimbutane 2011), and educational language policy in the Lao PDR (Cincotta-Segi 2009).

Hornberger and Johnson (2007) discuss the ethnography of language policy as a method that can illuminate and inform multiple types of language planning (status, corpus, and acquisition), illuminate and inform language policy processes (creation, interpretation, and appropriation), marry a critical approach with a focus on educator agency, and examine

connections across the various layers and levels of LPP activity. Johnson (2009) presents a heuristic for the scope of data collection in ethnography of language policy:

1. Agents—include both the creators of the policy and those responsible for policy interpretation and appropriation.
2. Goals—refers to the intentions of the policy as stated in the policy text.
3. Processes—creation, interpretation, and appropriation of policy text and discourse.
4. Discourses that engender and perpetuate the policy—the discourses within and without the policy; i.e., the discourses (whether explicit or implicit) within the language policy texts, intertextual and interdiscursive connections to other policy texts and discourses, and the discursive power of a particular policy.
5. The dynamic social and historical contexts in which the policy exists—an ethnography of language policy is interested in the dynamic social, historical, and physical contexts in which language policies are created, interpreted, and appropriated.

Sometimes separated from, but often combined with ethnographic and other qualitative research approaches, different forms of discourse analysis have been leveraged to illuminate language policy creation, interpretation, and negotiation. Traditions of discourse analysis have their own methodological history, representing different foci, questions, and epistemological orientations. For example, grounded in the work of Sacks and Schegloff (Sacks et al. 1974), conversation analysis (CA) focuses on the organised set of practices in the structure of human interaction. While CA attends closely to the interactional moves among conversation participants, critics have argued that the context in which the interaction occurs has not been a focus. Critical discourse analysis (CDA), on the other hand, places the social context, and particularly issues of power and inequality, as central to the analysis of the text. Scholars like Fairclough (2010) and Wodak (1996) have proposed analytic frameworks for making connections between texts and contexts, often establishing complex intertextual and interdiscursive paths between multiple levels of creation, interpretation, and recontextualisation.

Discourse analysts who ground their work in linguistic anthropology (e.g., Wortham 2005; Agha 2003) also consider the context to be of tantamount importance, but distinguish their work from CDA by an

in-depth, ethnographically-informed, understanding of the context (see Mortimer 2013). With a similar commitment to understanding local complexity, sociolinguistic ethnography (Pérez-Milans 2012) combines fine-grained discourse analysis with ethnographic data collection. Echoing Blommaert's criticism of CDA, Pérez-Milans (2012: 63) argues that other discourse analytic approaches "have emphasised stable macro-societal processes which are taken for granted and projected on the discourse as background facts". Instead, sociolinguistic ethnography attends to the links between local practices and macro-level social processes, neither of which are monolithic or static.

Each of these discourse analytic traditions has been applied to LPP studies. Bonacina's (2010) analysis of an induction classroom in France takes up Spolsky's (2004) inclusion of "language practices" as an essential part of language policy. Utilising CA, Bonacina examines how code-switching practices emerge as a norm of interaction within the classroom, forming what she calls a "practiced language policy". Cincotta-Segi's work (Cincotta-Segi 2009, 2011) on language policy and education in the Lao People's Democratic Republic (Lao PDR) reveals how impacts of a language policy cannot necessarily be predicted based solely on the language of that policy or the perceived intentions of its authors. She combines ethnography and CDA to illuminate how the de jure intentions of the state—as expressed through governmental policy and analysed within interviews with Ministry of Education officials—are marked by heterogeneity, and the interpretation and appropriation by teachers are equally multifarious. Cincotta-Segi's research reveals the importance of considering the multiple intentions, ideologies, and discourses that engender policy text and discourse, as well as the multiple, unpredictable, and agentive forms of interpretation and appropriation unique to a particular context. Her analysis balances a critical focus on the power of language policy as a mechanism of hegemony (through CDA) with an understanding of the power of language policy agents (through ethnography). Mortimer (2013: 67) draws upon linguistic anthropology to examine language policy as "a constellation of communicative events" in which linguistic signs and social meaning emerge and change. She incorporates speech chain analysis (Agha 2003) to trace a chain of communicative events that connect macro-level language policy with local educational practice in Paraguay. Mortimer traces two distinct meanings of what it means to speak Guaraní across diverse language policy texts, talk, and practices.

6   Current Debates

6.1   *Structure and Agency*

An essential challenge facing LPP researchers is how to balance critical analyses of the power of policy (discourses) with an empirical understanding of the agency of policy actors. The conceptualisation of, and relationship between, structure and agency, is a central concern in social theory and these concepts are taken up by other disciplines including Public Health (Karlsen and Nazroo 2002), Education (Archer 2002), Business (Gorton 2000), and environmental studies (McLaughlin and Dietz 2008). More recently, these concepts are featured in LPP research specifically, and applied and educational linguistics more generally. However, structure and agency are often underspecified or misidentified as competing forces within a reductive dichotomous relationship wherein language policy is portrayed as the hegemonic structure while agency is described as a reaction to that structure, which takes the form of individual free will. Furthermore, agency is often used as a synonym for resistance.

Presenting a more robust analysis, the application of structure and agency to LPP is featured in two collections of research—an edited volume by Glasgow and Bouchard (2019), and a special edition of *CILP* edited by Liddicoat and Taylor-Leech (2021). In their introduction, Glasgow and Bouchard (2019) propose a solution for balancing the tension between a theoretical focus on policy as structure/discourse and the emphasis on agency in the interpretation and appropriation of language policies. They argue that structure and agency do not have a dichotomous relationship but are mutually constitutive, dialectal, and dialogic and we "can discern echoes of each in both" (p. 49). A central tenet is the rejection of agency as (only) free will or unfettered decision making by language policy agents—even if free will is still a possibility—and a re-centring of agency as a process that is impacted by educational and economic inequality. A similar point is made by Blommaert (2005) who argues that conceptualisations of agency should not eliminate human creativity but situate it "in a wider frame of constraint…in the borderline zone of exiting hegemonies" (p. 99, 106). Similarly, in the introduction to the special issue of *CILP*, Liddicoat and Taylor-Leech (2021) review competing conceptualizations of structure and agency, including the critical realist perspective, which proposes that structure and agency

are distinct phenomena, and the constructivist perspective, which is influenced by Giddens (1979) and emphasizes the duality of structure and agency. Liddicoat and Taylor-Leech point out that LPP studies do not always make clear the way in which structure and agency are understood and used in the research.

Other influential publications that have considered the role of structure and agency within LPP processes include Menken and García's (2010) edited volume *Negotiating Language Policies in Schools: Educators as Policymakers*, which prioritises this debate and, in their introduction, García and Menken (2010) argue that teachers are the final arbiters of language policy implementation. Expanding on this argument, Johnson and Johnson (2015) propose that *language policy arbiters* are individuals who wield a disproportionate amount of LPP power relative to other individuals in a particular context. They argue that the heterogeneity of language policy texts, and the diversity of the sociolinguistic and sociocultural contexts in which policies are interpreted and appropriated, creates opportunities for creative and unpredictable forms of interpretation and appropriation. Nevertheless, while emphasising the relative power of human agency, they also argue that some individuals get positioned as more powerful and this positioning tends to rely on traditional/dominant sociolinguistic and sociocultural hierarchies (Barakos 2016; Cincotta-Segi 2009; Johnson 2012).

It is tempting to equate structure with macro-level social processes/systems and agency with micro-level human behaviour, yet both macro- and micro-discourses, and both structure and agency, can emerge in a single discursive event or shape a single policy document. Policy texts, discourses, and practices are heterogeneous, and ideologies are multiply layered, and all can change from context to context over time. Discourse analytic techniques empirically uncover how LPP processes can lead to both social change (Mortimer 2013) and hegemony (Savski 2020). They contribute to a theory of social change within language policy processes, thus helping to develop critical theories for sociolinguistics more generally. They complicate well-established definitions and conceptualisations, including the time-honoured macro-micro dialectic. It was necessary to borrow from other disciplines in the past, but a new wave of LPP research will help create theories and methods

specific to the field and discursive approaches give the "critical" in Critical Language Policy some methodological teeth and empirical rendering (Barakos and Unger 2016).

In LPP research, while structure and agency have been usefully combined with other theories within the field, the application of agency and structure in LPP research has tended to appropriate the dichotomous view of these concepts. In order for the potential of structure and agency to both further research and provide solutions and strategies in creating and implementing language policy, a more dynamic understanding of these concepts needs to be explored. In a forthcoming chapter, we outline five findings for LPP scholars to consider when applying these concepts (Johnson and Vernon, forthcoming): (1) structures are not necessarily immutable and research should consider both stability and change; (2) structures are both synchronic and diachronic and, therefore, research should consider the historical processes that engendered current social structures; (3) structure is both constraining and enabling and research should consider how language policy both normalises particular marginalising linguistic and sociolinguistic norms and sometimes champions multilingualism as a resource; (4) agency is not just individual free will, it includes collective activities; and (5) neither structure nor agency is necessarily intentional and can emerge by accident.

6.2   *Macro-Micro Dialectic*

There is general agreement in the field that language policy should be conceptualised and studied as multiply levelled or layered. Yet, criticism has focused on monolithic and static depictions of the relationship between macro and micro because they do not account for the multiple constraints on social interaction that can change over time. Critics argue that research focusing on the micro tends to overestimate the individual's ability for novel and seminal action (Wortham and Reyes 2015). Questioning and re-conceptualising the macro-micro dialectic is becoming an important feature within LPP research. While a layered understanding of context is implicit in the macro-micro distinction (especially when other "meso" layers are added), ideologies are also layered and can change. As Mortimer and Wortham (2015: 163) argue, "[I]nstead of connecting micro-level events to macro-level structures (e.g. connecting a classroom

language practice to an official policy), we must explore heterogeneous domains and scales of social organisation relevant to understanding meaningful social action". Therefore, as Hult (2010) has argued, within any discursive event, there are many potential sociolinguistic scales at work, and the analyst identifies how the unique configuration of semiotic resources is made relevant within the interaction. Nevertheless, the benefit of replacing levels/layers with 'scales' is still being debated, especially in LPP research that must characterise physical contexts (as opposed to, or as well as, discursive contexts) and/or does not make use of discourse analytic techniques.

Menken (2008) uses the term "arbiter" to characterise the power of teachers as the ultimate decision makers in how a policy is implemented. Johnson and Johnson (2015) expand on this notion and describe all individuals with potentially powerful influence on the language policy process as language policy arbiters. While LPP is a multi-layered process, and teachers may be the ultimate arbiters in classroom implementation of policy, language policy power is differentially allocated across, and within, educational institutions, contexts, and layers of language policy activity. For example, Johnson (2012) reveals how a change in one U.S. school district leadership position led to a drastic change in language policy for the entire district, transitioning from a focus on the value of bilingualism to an emphasis on acquisition of English. The power of the language policy arbiter is such that they have a singular impact on educational activities within the institutions and contexts they have contact with.

6.3   *Researcher Positionality*

Another important trend in LPP research is the focus on researcher positionality and subjectivity. While the term is less often defined, it has become accepted that language policy scholars are "critical" since so much of the research is focused on power and/or committed to social justice. Yet "activism" and "research" are often separated in ways that perpetuate divisions between participants and observers and reify objectivist epistemologies. De Costa (2014) argues that scholars should not only describe, critique, and transform discriminatory practices and social inequality, they should also maintain critical reflexivity regarding the position of the researcher. Similarly, in Lin's (2015) interrogation of researcher epistemology in LPP, she argues that a critical perspective

combines self-reflection of the researcher's position in institutional hierarchies with an interrogation of how such institutional hierarchies produce and reproduce domination and subordination. According to Lin (2015: 26), the goal of the researcher should be knowledge co-construction: "In the critical research paradigm, both the researcher and the researched are subjects of knowing and enter into a dialogue on equal footings". Nevertheless, Lin highlights a historical weakness in critical research, which has impacted policy less often than positivist studies. The solution might be a "dialogue with different parties in each situated context in LPP" (p. 30), which recognises the strength in collaboration and acknowledges the limited nature of a single methodological position.

One of the areas in which researchers can openly articulate their own positions and subjectivities is in the written report. Ramanathan (2006, 2011) compellingly grapples with her own ideological conflict in researching-texting practices and criticises LPP scholarship that presents its textual products in hermetically sealed ways—"a modus operandi that often leaves little room for addressing uncertainties and tensions in the researching-texting process" (Ramanathan 2011: 256). Instead, she emphasises that how we develop meaning in the texts we create is a process that should be rendered "porous, unstable, and changeable" (p. 247). Echoing Fishman's (1994) critique, she argues that it is crucial to interrogate the ideological aspects of research as a social practice, while also permitting a way to "complexify the researcher's voice" (p. 268). Similarly, Canagarajah and Stanley (2015: 41) argue genre conventions in academic writing make it challenging to give voice to minority communities, yet it is essential that LPP scholars push back against positivistic writing genres that attempt to synthesise research findings into generalisable and monolithic "truths": "Since the subjects exist in the report only through the voice of the researcher, there is a tendency for their complexity to be suppressed and their identity to be generalised (or essentialized)".

Consideration of ethics and advocacy in sociolinguistic research is nothing new. For example, Labov (1982) and Rickford (1999) have grappled with the responsibility that language scholars have to the speech communities in which they collect data and both have actively promoted social justice efforts in courts, legislatures, and public discourse. Nevertheless, Labov has consistently separated research from advocacy. A new wave of LPP scholars is considering how to combine research with application, observation with action, and critique with advocacy. There

is a growing sentiment, absent in earlier language planning research, that LPP scholars should also be active in language policy efforts, and these two activities—research and action—can be combined in theoretically and methodologically robust ways, despite objectivist arguments to the contrary. Admittedly, this type of engagement may prove more compelling for qualitative researchers than those who specialise in surveys, economics, or demography. Furthermore, this type of engagement may imperil the position of LPP scholars who advise and consult on language planning activities, and need to retain the credibility imparted by scientific objectivity. Indeed, what is promoted here is a transdisciplinary approach that combines diverse theories, methods, and epistemologies to promote social justice in LPP documents, processes, and impacts.

7   CRITICAL EMPIRICAL APPROACHES AND SOCIAL JUSTICE: CONCLUDING THOUGHTS

By criticising LPP processes—and the institutions responsible for their enactment (like schools)—CLP has highlighted the continued minoritisation of historically marginalised people in schools and society. Kids who speak a non-dominant language or dialect, for example, still lack equal educational opportunity in schools, which is often the result of a combination of racism and language policy, emanating from what Rosa and Flores have referred to as raciolinguistic ideologies (Flores and Rosa 2015). Yet, in foregrounding the role of social justice in LPP processes, LPP research has also emphasised how LPP processes in schools and society can promote ideological and implementational spaces for historically marginalised languages and speakers of those languages. The potentially emancipatory impact is an essential contribution from Nancy Hornberger and Joshua Fishman, and the other authors in this book. What is clear is a balanced perspective and method are needed—which considers both the hegemonic impact and disempowering reality of LPP processes and the emancipatory possibilities. This requires us to move beyond macro-level critique to fieldwork and empirical action. This requires a transdisciplinary approach that transcends methods and techniques tethered to specific disciplines to solve LPP problems. Scholars from diverse epistemological, methodological, and theoretical backgrounds—linguistics, economists, educationists, political theorists, etc.—will all be needed.

References

Agha, A. (2003). The social life of cultural value. *Language and Communication* 23, pp. 231–273. https://doi.org/10.1016/S0271-5309(03)00012-0

Alexander, N. (1989). *Language Policy and National Unity in South Africa/Azania*. Cape Town: Buchu Books.

Archer, M. (2000). *Being Human: The Problem of Agency*. Cambridge: Cambridge University Press.

Archer, R. (2002). *Education policy and realist social theory: Primary teachers, child-centered philosophy and the new managerialism*. Routledge.

Barakos, E. (2016). Language policy and governmentality in businesses in Wales: A continuum of empowerment and regulation. *Multilingua*, 34 (4), pp. 361–391. https://doi.org/10.1515/multi-2015-0007

Barakos, E., and Unger, J. W. (Eds.). (2016). *Discursive approaches to language policy*. Palgrave Macmillan.

Blommaert, J. (2005). *Discourse: Key Topics in Sociolinguistics*. Cambridge: Cambridge University Press.

Bonacina, F. (2010). A conversation analytic approach to practiced language policies: The example of an induction classroom for newly-arrived immigrant children in France. Ph.D. Thesis. The University of Edinburgh.

Bouchard, J., and Glasgow, G. P. (eds.) (2019). *Agency in Language Policy and Planning: Critical Inquiries*. London: Routledge.

Canagarajah, S., and Stanley, P. (2015). Ethical considerations in language policy research. In F. M. Hult and D. C. Johnson (eds.), *Research Methods in Language Policy and Planning: A Practical Guide*. Malden: Wiley Blackwell.

Chimbutane, F. (2011). *Rethinking Bilingual Education in Postcolonial Contexts*. Clevedon: Multilingual Matters.

Cincotta-Segi, A. (2009). 'The big ones swallow the small ones.' Or do they? The language policy and practice of ethnic minority education in the Lao PDR: A case study from Nalae. Ph.D. Thesis. Canberra: The Australian National University.

Cincotta-Segi, A. (2011). 'The big ones swallow the small ones'. Or do they? Language-in-education policy and ethnic minority education in the Lao PDR. *Journal of Multilingual and Multicultural Development*, 32 (1), pp. 1–15. https://doi.org/10.1080/01434632.2010.527343

Cobarrubias, J. (1983). Ethical issues in status planning. In J. Cobarrubias and J. A. Fishman (eds.), *Progress in Language Planning: International Perspectives*, pp. 41–86. Berlin: Walter De Gruyter & Co.

Cooper, R. L. (1989). *Language Planning and Social Change*. New York: Cambridge University Press.

Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process*. London: Sage.

Davis, K. (1994). *Language planning in multilingual contexts: Policies, communities, and schools in Luxembourg*. Philadelphia, PA: John Benjamins.

Davis, K. A. (1999). Dynamics of indigenous language maintenance. In T. Huebner and K. A. Davis (eds.), *Sociopolitical Perspectives on Language Policy and Planning in the USA*, pp. 67–98. Amsterdam and Philadelphia: John Benjamins Publishing Company.

De Costa, P. (2014). Making ethical decisions in an ethnographic study. *TESOL Quarterly*, 48 (2), pp. 413–422. https://doi.org/10.1002/tesq.163

Fairclough, N. (2010). *Critical Discourse Analysis: The Critical Study of Language*. 2nd Edition. Harlow, England: Pearson.

Fishman, J. A. (1994). Critiques of language planning: A minority languages perspective. *Journal of Multilingual & Multicultural Development*, 15 (2–3), pp. 91–99. https://doi.org/10.1080/01434632.1994.9994559

Fishman, J., Ferguson, C. A., and Das Gupta, J. (eds.) (1968). *Language Problems of Developing Nations*. New York: Wiley.

Flores, N., and Rosa, J. (2015). Undoing appropriateness: Raciolinguistic ideologies and language diversity in education. *Harvard Educational Review*, 85 (2), pp. 149–171.

Foucault, M. (1978). *The History of Sexuality*. New York: Random House.

Foucault, M. (1982). The subject and power. *Critical Inquiry,* 8 (4), pp. 777–795.

Fowler, R., Hodge, B., Kress, G., and Trew, T. (1979). *Language and Control*. London: Routlege and Kegan Paul.

Freeman, R. D. (1998). *Bilingual Education and Social Change*. Clevedon: Multilingual Matters.

García, O., and Menken, K. (2010). Stirring the onion: Educators and the dynamics of language Education policies (looking ahead). In K. Menken and O. García (eds.), *Negotiating Language Policies in Schools: Educators as Policymakers*, pp. 249–261. London and New York: Routledge.

Giddens, A. (1979). *Central Problems in Social Theory: Action, Structure and Contradiction in Social Analysis*. Berkeley: University of California Press.

Glesne, C. (2016). *Becoming Qualitative Researchers: An Introduction*. New York: Pearson.

Gorton, M. (2000). Overcoming the structure–agency divide in small business research. *International Journal of Entrepreneurial Behavior & Research*, 6 (5), pp. 276–292.

Habermas, J. (1973). *Theory and Practice*. Boston, MA: Beacon Press.

Halliday, M. A. K. (1978). *Language as Social Semiotic*. London: Edward Arnold.

Haugen, E. (1972). The ecology of language. In A. S. Dil (ed.), *The Ecology of Language: Essays by Einar Haugen*, pp. 325–339. Stanford, CA: Stanford University Press.

Hill, R., and May, S. (2011). Exploring biliteracy in Māori-medium education: An ethnographic perspective. In T. L. McCarty (ed.), *Ethnography and Language Policy*, pp. 161–183. New York: Routledge.

Hornberger, N. H. (1988). *Bilingual Education and Language Maintenance*. Dordrecht, Holland: Foris Publications.

Hornberger, N. H., and Johnson, D. C. (2007). Slicing the onion ethnographically: Layers and spaces in multilingual language education policy and practice. *TESOL Quarterly*, 41 (3), pp. 509–532. https://doi.org/10.1002/j.1545-7249.2007.tb00083.x

Hult, F. M. (2007). *Multilingual Language Policy and English Language Teaching in Sweden*. Ph.D. Thesis. Philadelphia: University of Pennsylvania.

Hult, F. M. (2010). Analysis of language policy discourses across the scales of space and time. *International Journal of the Sociology of Language*, 202, pp. 7–24. https://doi.org/10.1515/ijsl.2010.011

Hymes, D. (1964). Introduction: Toward ethnographies of communication. *American Anthropologist*, 66 (6), pp. 1–34.

Johnson, D. C. (2009). Ethnography of language policy. *Language Policy*, 8 (2), pp. 139–159. https://doi.org/10.1007/s10993-009-9136-9

Johnson, D. C. (2012). Positioning the language policy arbiter: Governmentality and footing in the School District of Philadelphia. In J.W. Tollefson (ed.), *Language Policies in Education: Critical Issues* (2nd Edition). London and New York: Routledge.

Johnson, D. C., & Johnson, E. J. (2015). Power and agency in language policy appropriation. *Language Policy*, 14 (3), pp. 221–243.

Johnson, D. C., and Ricento, T. (2013). Conceptual and theoretical perspectives in language policy and planning: Situating the ethnography of language policy. *International Journal of the Sociology of Language*, 219, pp. 7–21. https://doi.org/10.1515/ijsl-2013-0002

Johnson, D. C., and Vernon, A. (forthcoming). Structure and agency in educational language policy. In F. Hult and B. Spolsky (eds.), *Handbook of Educational Linguistics*. Malden: Wiley.

Kaplan, R. B., Baldauf Jr., R. B., Liddicoat, A., Bryant, P., Barbaux, M.-T., and Pütz, M. (2000). Current issues in language planning. *Current Issues in Language Planning*, 1 (1), pp. 1–10. https://doi.org/10.1080/14664200008668003

Karlsen, S., and Nazroo, J. Y. (2002). Agency and structure: The impact of ethnic identity and racism on the health of ethnic minority people. *Sociology of Health and Illness*, 24 (1), pp. 1–20.

King, K. A. (2001). *Language Revitalization Processes and Prospects: Quichua in the Ecuadorian Andes*. Clevedon: Multilingual Matters.

Labov, W. (1982). Objectivity and commitment in linguistic science: The case of the Black English trial in Ann Arbor. *Language in Society*, 11 (2), pp. 165–201.

Leibowitz, A. H. (1984). The official character of language in the United States: Literacy requirements for immigration, citizenship, and entrance into American life. *Aztlan*, 15 (1), pp. 25–70.

Levinson, B. A. U., and Sutton, M. (2001). Introduction: Policy as/in practice—A sociocultural approach to the study of educational policy. In B. A. U. Levinson and M. Sutton (eds.), *Policy as Practice: Toward a Sociocultural Analysis of Educational Policy*, pp. 1–22. London: Ablex Publishing.

Liddicoat, A. J., and Taylor-Leech, K. (2021). Agency in language planning and policy. *Current Issues in Language Planning*, 22 (1–2), pp. 1–18. https://doi.org/10.1080/14664208.2020.1791533

Lin, A. M. Y. (2015). Researcher positionality. In F. M. Hult and D. C. Johnson (eds.), *Research Methods in Language Policy and Planning: A Practical Guide*. Malden: Wiley.

Madison, D. S. (2012). *Critical Ethnography: Method, Ethics, and Performance*. Thousand Oaks: Sage.

McCarty, T. (2002). *A Place to be Navajo: Rough Rock and the Struggle for Self-Determination in Indigenous Schooling*. London: Routledge.

McLaughlin, P., and Dietz, T. (2008). Structure, agency and environment: Toward an integrated perspective on vulnerability. *Global Environmental Change*, 18 (1), pp. 99–111.

Menken, K. (2008). *English Learners Left Behind: Standardized Testing as Language Policy*. Clevedon, UK: Multilingual Matters.

Menken, K., and García, O. (eds.) (2010). *Negotiating Language Policies in Schools: Educators as Policymakers*. New York: Routledge.

Mill, J. S. (1866). *August Comte and Positivism*. Cambridge: Cambridge University Press.

Mortimer, K. (2013). Communicative event chains in an ethnography of Paraguayan language policy. *International Journal of the Sociology of Language*, 219, pp. 67–99. https://doi.org/10.1515/ijsl-2013-0005

Mortimer, K., and Wortham, S. (2015). Analyzing language policy and social identification across Heterogeneous scales. *Annual Review of Applied Linguistics*, 35, pp. 160–172. https://doi.org/10.1017/S0267190514000269

Pérez-Milans, M. (2012). Beyond 'safe-talk': Institutionalization and agency in China's English language education. *Linguistics and Education*, 23, pp. 62–76. https://doi.org/10.1016/j.linged.2011.11.001

Proietti, M., Pickston, A., Graffitti, F., Barrow, P., Kundys, D., Branciard, C., Ringbauer, M., and Fedrizzi, A. (2019). Experimental test of local observer independence. *Science Advances*, 5, pp. 1–6. https://doi.org/10.1126/sciadv.aaw9832

Ramanathan, V. (2006). Of texts and translations and rhizomes: Postcolonial anxieties and deracinations and knowledge constructions. *Critical Inquiry in Language Studies*, 3 (4), pp. 223–244. https://doi.org/10.1207/s15427587 clis0304_2

Ramanathan, V. (2011). Researching-texting tensions in qualitative research: Ethics in and around textual fidelity, selectivity, and translations. In. T. McCarty (ed.), *Ethnography and Language Policy*, pp. 255–270. New York and London: Routledge.

Ricento, T. (2000). Historical and theoretical perspectives in language policy and planning. *Journal of Sociolinguistics*, 4 (2), pp. 196–213. https://doi.org/10.1111/1467-9481.00111

Ricento, T., and Hornberger, N. H. (1996). Unpeeling the onion: Language planning and policy and the ELT professional. *TESOL Quarterly*, 30 (3), pp. 401–427. https://doi.org/10.2307/3587691

Rickford, J. R. (1999). The Ebonics controversy in my backyard: A sociolinguist's experiences and reflections. *Journal of Sociolinguistics*, 3 (2), pp. 267–275.

Roman, L. G. (1993). Double exposure: The politics of feminist materialist ethnography. *Educational Theory*, 43 (3), pp. 279–308. https://doi.org/10.1111/j.1741-5446.1993.00279.x

Rubin, J. and Jernudd, B. H. (eds.) (1971). *Can Language be Planned? Sociolinguistic Theory and Practice for Developing Nations*. Honolulu: The University Press of Hawaii.

Rubin, J., Jernudd, B. H., Das Gupta, J., Fishman, J. A., and Ferguson, C. A. (1977). *Language Planning Processes*. The Hague: Mouton Publishers.

Ruiz, R. (1984). Orientations in language planning. *NABE Journal*, 8 (2), pp. 15–34. https://doi.org/10.1080/08855072.1984.10668464

Sacks, H., Schegloff, E. A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50 (4), pp. 696–735. https://doi.org/10.2307/412243

Savski, K. (2020). Local problems and a global solution: Examining the recontextualization of CEFR in Thai and Malaysian language policies. *Language Policy*, 19 (4), pp. 527–547. https://doi.org/10.1007/s10993-019-09539-8

Spolsky, B. (2004). *Language Policy*. Cambridge: Cambridge University Press.

Spolsky and Shohamy. (2002). Editorial. *Language Policy*, 1, pp. 1–2.

Tollefson, J. W. (1991). *Planning Language, Planning Inequality: Language Policy in the Community*. London: Longman.

Tollefson, J. W. (2002). Introduction: Critical issues in educational language policy. In J. W. Tollefson (ed.), *Language Policies in Education: Critical Issues*. Mahwah, NJ: Lawrence Erlbaum Publishers.

Tollefson, J. W. (2006). Critical theory in language policy. In T. Ricento (ed.), *An Introduction to Language Policy: Theory and Method*, pp. 42–59. Malden, MA: Blackwell Publishing.

Tollefson, J. W. (2012). Language policy in a time of crisis and transformation. In J. W. Tollefson (ed.), *Language Policies in Education: Critical Issues* (2nd Edition), pp. 11–32. London and New York: Routledge.

Wiley, T. (1999). Comparative historical analysis of U.S. language policy and language planning: Extending the foundations. In T. Huebner and K. A. Davis (eds.), *Sociopolitical Perspectives on Language Policy and Planning in the USA*. Philadelphia: John Benjamins Publishing Company.

Wodak, R. (1996). *Disorders of Discourse*. London: Longman.

Wortham, S. (2005). Socialization beyond the speech event. *Journal of Linguistic Anthropology*, 15 (1), pp. 95–112.

Wortham, S., and Reyes, A. (2015). *Discourse Analysis Beyond the Speech Event*. New York: Routledge.

Language Policy as Public Policy

Michele Gazzola

Abstract This chapter presents and discusses the public policy approach to language policy and planning (LPP). The chapter critically examines the idea according to which everyone (including single individuals) can make language policy, and it clarifies the differences between (individual) language practices and (public) language policy. It argues that government (at different levels, from the local to the national, and even supranational) is the central, albeit of course not exclusive, agent in the language policy process. Language policy is therefore presented as a particular form of public policy that can be examined using the policy cycle model, which is the standard analytical framework in public policy studies. The model comprises several stages, namely the emergence of a language issue in society, followed by agenda-setting, policy formulation and adoption, implementation, and finally evaluation. The chapter presents these five stages of the cycle and the relationships between them and clarifies how the various disciplines involved in LPP can contribute to their study. It presents some concepts that are central in the design

---

41

M. Gazzola et al., *Epistemological and Theoretical Foundations in Language Policy and Planning*, https://doi.org/10.1007/978-3-031-22315-0_3

of language policies, such as programme theory, policy instruments, and indicators. The chapter also discusses the limits of the model. It emphasises the political and pragmatic aspects of LPP, and the need to engage with decision makers' practical concerns. In this sense, it contributes to restoring the theoretical and epistemological links between LPP research and public policy studies that have gradually weakened after the 'critical turn' in LPP.

1   Introduction

This chapter presents and discusses the public policy approach to language policy and planning (LPP). It reviews some central concepts in public policy studies, such as policy cycle, programme theory, policy instruments, and indicators, and it explains how they can be applied to LPP. Although this contribution has been written primarily for an audience of applied linguists, sociolinguists, and decision-makers, it can also be useful for students in public policy and administration, political science, economics, and political philosophy who are interested in approaching LPP from this perspective.

This contribution starts from the observation that language policy can be studied like other well-established forms of public policy/social policy in areas such as the environment, income support, transportation, education, or healthcare. From this perspective, language policy can be defined as a public policy aimed at addressing a social, economic, political or organisational issue related to the management of linguistic diversity in a given territory. Language policy, therefore, is a response by the government to an issue politically defined as collective in nature involving language or languages. While language policy obviously deals with language(s), it is "ultimately oriented towards non-linguistic ends" (Cooper 1989: 35). For example, the target of language policies protecting and promoting minority languages is not only language as such, but also (and more importantly) the community of speakers and

the protection and promotion of their political and cultural rights to use the language in social life.

This chapter contributes to restoring the theoretical and epistemological links between LPP research and public policy studies, which have gradually weakened after the 'critical and ethnographic turns' in LPP (see Johnson, this volume). Although the importance of the public policy approach in LPP was already clear from the origins of the discipline (Rubin and Jernudd 1971), still too little attention in research has been paid to the operational and organisational aspects of LPP, including "the development, implementation, and evaluation of specific language policies" (Ricento 2005: 18). As Leigh Oakes observes,

> the critical and ethnographic turns have undeniably allowed LPP to make important theoretical and methodological advances. Nonetheless, as LPP has matured as a field of academic inquiry with its own intellectual objectives and specialised vocabulary, it has arguably also drifted further away from the concerns of policymakers needing to make and justify choices about language. In order to help realise the transformative aspirations expressed especially by critical approaches to LPP, the field would do well to find ways to reconnect with these groups, to decrease, not increase, the gap between research and practice already noted in the pioneer years. The way forward lies perhaps not in a new *turn* so much as a *re*turn; or more specifically, a rediscovery of the field's more pragmatic and outward-looking origins. (Oakes 2023, forthcoming)

This chapter argues that to promote this 'return' to the pragmatic origins of LPP research, we do not need to adopt again the old approaches and theories of LPP from the 1960s–1970s (often defined as 'technocratic' by critical LPP scholars). Rather, it means to re-establish and deepen the epistemological and theoretical links with contemporary public policy studies.

As already explained in the introduction to this book, language policy affects the structure (or corpus), the functions (or status), and the acquisition of a language. This chapter focuses on the last two aspects of language planning (see Gobbo, this volume, for corpus planning), and it is organised as follows: Section 2 clarifies the epistemological and theoretical implications of studying language policy as a public policy. Section 3 presents the fundamental framework used in public policy studies to study public policies, that is, the policy cycle model, and it explains the extent to which the different disciplines involved in LPP can contribute to the understanding of each stage of the cycle. Section 4 completes the discussion by examining the relationships between macro-level and meso-level

language policies, as well as the relationships between LPP and other types of public policy. The last section concludes the chapter.

2   PUBLIC POLICY AND LANGUAGE PRACTICES

Knoepfel et al. define a public policy 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" (2007: 24). There are, of course, many alternative definitions of public policy in the literature, and it would be tedious to report them here. Instead, it is more interesting to concentrate on some common elements of the various definitions of public policy (see Howlett et al. 2020: 1–19).

Firstly, the primary agent of public policy is the *government*, at different institutional levels, since it has the ability and the legitimacy to make authoritative collective decisions on behalf of citizens. This, of course, does not exclude private actors from playing a role in the policy process (see Sect. 3). It is important to recall that the term 'government', in its broadest sense, refers to the institutional processes through which collective (and usually binding) decisions are made, and it is not necessarily restricted to 'the government of the day' in the sense of 'the executive' (Heywood 2019). Decisions made by local authorities, e.g., in the field of education, are also a relevant object of public policy studies. While in the early days of the discipline "LPP was understood primarily as the plans and policies formulated and implemented at the national level by ministries of education and similar state authorities" (Tollefson and Pérez Milans 2018: 7), more recent contributions include a broader range of institutional levels (from the local to the international), as well as several policy areas in which governments operate. Education is the obvious example, but there are many other areas such as healthcare, the integration of migrants into the labour market, the organisation of the public administration in officially multilingual countries, and the international protection of intellectual property rights. Language policy, in some cases, is just the linguistic component of a broader social policy (more on this in the conclusions of this book).

Secondly, policymaking is about making choices to do something or nothing to address a public problem/issue. If abstaining from doing something is a deliberate decision, then also 'doing nothing' is a public

policy. One could notice some similarities between this aspect of public policy and earlier definitions of LPP as attempts to solve 'language problems' (see the introduction to this volume). As Johnson notes "early language planning research was dominated by objectivism. There were language *problems* to be solved … and solutions were proposed in the form of language planning frameworks, models, and conceptualizations" (Johnson 2018: 53). Solutions, of course, were proposed by 'experts'. This approach is seen as 'technocratic' in critical LPP, because early LPP was influenced by an interest in excluding sociopolitical variables in the objective science of language planning (Johnson, this volume). This is not, however, the approach followed in public policy studies. As shown in more detail in Sect. 4, the 'public problems' or issues to be addressed through public policy emerge precisely in and from the public debate, that is, in society, and they are subsequently constructed and framed in the political debate in the agenda-setting phase *before* the formulation of possible solutions takes place. As Rossi et al. note in their book *Evaluation*, "social problems are not objective phenomena. Rather, they are social constructions involving assertions that certain conditions constitute problems that require public attention and deliberate, organized intervention" (Rossi et al. 2019: 34).

Thirdly, the content of a policy is made of a selection of *goals* and the *means* to achieve them. Public policy, therefore, is *goal-oriented*. This means that establishing logical connections between means and goals is an intrinsic component of policy design in any policy area, and not a specific feature of 'technocratic' approaches. Goals, however, are not technical—not necessarily at least (e.g., building a bridge)—since they are defined as a result of a political process. Finally, decisions are always made under constraints, and these constraints can be either technical (e.g., feasibility, state of knowledge) or political (e.g., norms, values and ideas). The level of constraints determines the capacity of the government to implement policies.

While in classic LPP, as shown in the introduction to this book, language policy was clearly viewed as a type of public policy designed and implemented directly or indirectly by government, many of the following definitions of language policy encompass virtually any actor in society who, at any level, tries to influence the linguistic behaviour of other people. Kaplan and Baldauf, for example, argue that the term language planning encompasses "everything from government macro-level national planning to group or individual micro-level planning"

(1997: 27), including "unplanned" forms of language planning (Baldauf 1994).

One of the most important consequences of this shift has been that the distinction between policy and practices has gradually faded away. In his influential model, Spolsky (2012) argues that language policy is made up of three main components. The first is the 'language practices' of the members of the speech community, that is, the habitual pattern of selecting among the varieties that make up its linguistic repertoire (i.e., words, sounds, grammatical choices, as well as choices of registers). The second element is 'beliefs' about language and language use, that is, ideas about what appropriate language practices are. In other words, language beliefs are "language policy with the manager left out, what people think should be done" (Spolsky 2004: 14), while language practices are what people actually do. The third component of language policy in Spolsky's definition is 'language management', that is, the efforts by *some* members of a speech community who have, or believe they have, authority over the other members to modify their linguistic practices. Anyone can be a language manager, from legislative assemblies writing a national constitution down to a family member trying to persuade others in the family to speak a heritage language (Spolsky 2004: 8).

In a recently revised version of his theory, Spolsky (2019) introduces two components to the definition of language policy. The first one is 'language advocates'. Language advocates are individuals or groups who wish to alter other people's practices without having the authority of language managers. Finally, he incorporates 'self-management' in the model, that is, "attempts of speakers to modify their own linguistic proficiency and repertoire" (2019: 326). Spolsky's over-extended definition of language policy and its actors is consistent with Kaplan and Baldauf's approach presented above. In recent years, the term language policy has been used to describe individuals' choices as to which languages to use to communicate within the family (i.e., 'family language policy'), in particular with children (see Caldas 2012; King and Fogle 2017 for an overview).

The consequences of this gradual broadening of the definition of language policy, however, are not negligible. The semantic space of the term 'language policy' has been stretched to such an extent to embrace virtually anything people decide to do with languages. Therefore, it has become so vague as to decrease in usefulness, because it does not facilitate a clear differentiation between 'policy' proper and 'practices', and between the roles and decisions of different actors in the policy process.

It is well known that the term 'policy' in English (but not necessarily in other languages) has a broad meaning, that is, "a definite course or method of action selected from among alternatives and in light of given conditions to guide and determine present and future decisions" (Webster dictionary). In an organisational context, for example, 'policy' indicates the decisions and the guiding principles of the organisation in relation to a certain question. One can refer to the 'baggage policy' of an airline to refer to the decisions about the size and the weight of passengers' baggage, or the 'smoking policy' of a train company, or the 'privacy policy' of a website. In this sense, it is roughly equivalent to 'rules' or 'regulations'. This breadth of meaning, however, does not help in defining the boundaries of LPP with respect to other disciplines, in particular sociolinguistics. What are the differences and the boundaries between LPP and sociolinguistics if policies and practices are conflated? As Johnson (2013: 23) correctly asks, what is not language policy, then?

In the public policy approach to LPP, these differences are clearer. Individuals do not make 'policies'; they take decisions which result in *practices*, and these practices are the focus of study of sociolinguistics.[1] By contrast, a government's decisions involving languages result in *policies*, and these are precisely the object of the study of a public policy approach to LPP. For example, a man who decides to reduce his consumption of cigarettes from 20 per day to just one because members of an association for the prevention of lung cancer convinced him to do so is not changing his 'smoking policy'; he is just changing his habits and trying to quit smoking. Individuals' practices include not only their routine behaviour (e.g., smoking), but also what they may do to change that behaviour (e.g., to quit smoking). From a public policy approach, the decisions of a father and a mother as to what languages to use or not use with their children at home are viewed as practices, actual individuals' behaviour. Such individuals' practices are simply the result of conscious decisions at the individual level (as opposed to purely impulsive reactions, e.g., dictated by immediate danger or fear) that people who are able to reason make routinely in their lives. They are not comparable with an institutionalised political process that typically produces documents or guidelines and involves the action of government through a series of policy instruments.

---

[1] For traditional approaches to sociolinguistics, I refer to Ammon et al. (2004), and Bayley et al. (2013). Poststructural approaches to the study of 'language and society' are presented in García et al. (2017).

Influencing people's attitudes and behaviour is usually the *target* or the outcome of public policy, but it is not part of public policy itself. Clearly, there are strong relationships between LPP in the public policy approach and sociolinguistics (see Sect. 3). A sound understanding of individuals' practices is crucial in order to plan policy, and the outcomes of a public policy are evaluated precisely in terms of their capacity to affect the evolution of individuals' practices and representations. But the two things are epistemologically and theoretically different.

Spolsky's model, therefore, is certainly useful when studying the ecology of language in a community and characterising the role of various agents in influencing language change. It is not the most suitable, however, when examining public policies targeting languages. Studying LPP as a public policy means focusing on the action of the government in addressing language issues, identifying and explaining the different phases of the policy process and how these phases are related to each other. It also means examining how different types of actors influence the process that leads to government decisions, and what the impact is of non-discursive factors (e.g., the strength of the administrative apparatus of the state) on the policy process and on speakers' attitudes and practices. Moreover, it requires an understanding of what the determinants of a good policy design are, and this involves the study of the different policy instruments that were designed or developed for the purpose of implementing those policies (e.g., information campaign, financial incentives, and regulations), as well as the indicators used in monitoring language policy implementation. Finally, it requires the empirical study of the effects of policy on the evolution of people's linguistic practices and representations. This framework is better provided by the 'policy cycle' briefly presented in the next section. It enables the connection of the structure and actors through institutions, instruments, and the analysis of practices. A fully-fledged and more in-depth application of the policy cycle model to the study of LPP is contained in the various chapters of the *Routledge Handbook of Language Policy and Planning* (Gazzola, Grin et al. 2023, forthcoming), to which the reader can refer.

3   THE POLICY CYCLE MODEL

The policy cycle is the standard heuristic framework/model employed to characterise and examine the different steps of the policy process,

as well as the relationships between them.[2] The policy cycle model presents policies as sequential parts or stages that correspond to applied problem-solving. These steps are (1) emergence of a collective issue, (2) agenda-setting, (3) consideration of policy options, (4) decision-making, (5) implementation, and (6) evaluation. Evaluation provides feedback on what has been done and the results obtained. It can lead to the successful conclusion of a policy if the problem has been solved, to a redefinition of the public policy itself, or to its eventual abandonment. In short, evaluation can be the end point of a programme and at the same time a new starting point.

There are some differences in the literature about the exact definition of each phase. Some contributions merge steps 3 and 4 into one single phase called 'policy formulation and adoption' (Mintrom and Williams 2013; Knoepfel et al 2015), because, as Jann and Wegrich note, "a clear-cut separation between formulation and decision-making is very often impossible" (Jann and Wegrich 2007: 48). This is the choice made in this chapter too. In other contributions, the first phase and the second phase are merged into 'agenda-setting' (Howlett et al. 2020), while introducing a separation between public (or systemic) agenda, i.e., the set of issues debated in the public space, and institutional (or formal) agenda, which instead is the policy agenda of government in the strict sense. This gives the circular flow chart shown in Fig. 1 that will be named the 'language policy cycle'.[3]

In order to avoid misplaced criticisms, it is important to clarify how this model can be used in LPP research and in practical language planning. The model has two main functions. First, it can be used by researchers to describe and analyse the language policy process, because it offers a systematic view of the different stages in the process leading from public and political debate to the design, implementation, and evaluation of language policy. The policy cycle framework "facilitates an understanding of a multidimensional process by disaggregating the complexity of the process into any number of stages and sub-stages, each of which can be

---

[2] Jann and Wegrich (2007) review the most important criticisms to the model and the responses to these criticisms.

[3] The term 'language policy cycle' is used also in Canagarajah (2005: 157–158) to denote "the different stages of language planning, that is, before, during, and after implementation". Canagarajah, therefore, adopts a generic definition of cycle which does not refer to the policy cycle model.
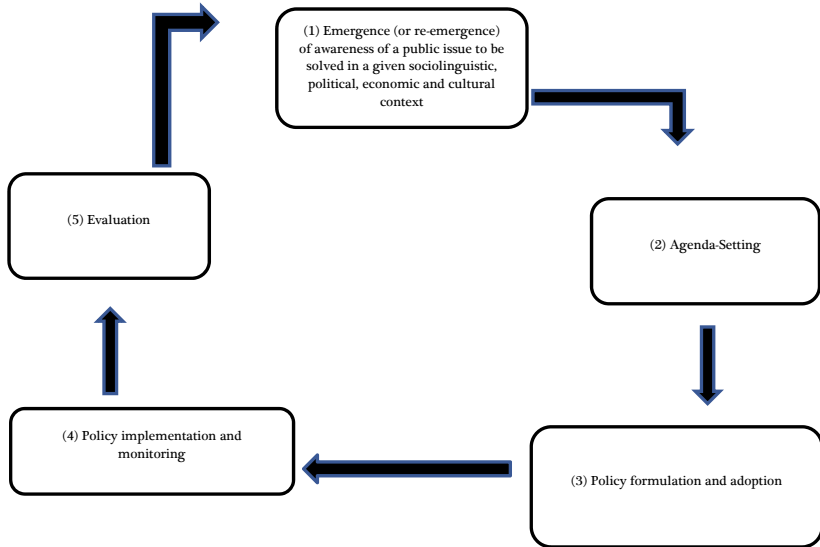
Fig. 1   Language policy cycle

investigated alone or in terms of its relationship to any or all the other stages of the cycle" (Howlett et al. 2020: 12). In addition, it is versatile, because it can be used at different levels of policymaking, from the local to the international, and it facilitates comparative analyses. Finally, it is inclusive, in the sense that it permits the examination of the intertwined roles of different actors, ideas, and organisations involved in language policy formulation, and not just governmental units formally charged with the task.

It is important to emphasise that the model does not have the ambition to reflect actual policymaking precisely. Policymakers do not always address and solve public problems in this systematic and linear way; they can skip some steps. For example, some decisions are never implemented, and some choices may be made before the debate or the agenda-setting phase. What is presented in the policy cycle, therefore, is an abstract and ideal version of the policy process. The model does not prescribe that all phases should and will be followed, but simply that the policy process typically involves different stages, and that these stages can be studied on

their own or in relation to other stages. In sum, the model is useful to organise knowledge about the policy process, its actors, and outcomes.

The second function of the model is to assist policymakers in actual language planning, by clarifying the phases and the challenges of the policy process. From this perspective, the model can be used as a practical guide, in particular in the stages of policy design, implementation, and evaluation. Recent examples of the application of this framework include UNESCO's "Global action plan of the International Decade of Indigenous Languages 2022–2032" (UNESCO 2021), and the "General Language Policy Plan for the Friulian Language 2021–2025" of the Autonomous Region Friuli-Venezia Giulia in Italy (ARLeF 2021).

3.1 Emergence of a Language Issue

The emergence and awareness of a 'language problem' in society is the preliminary stage of the cycle. Today, we would say 'issue' or 'question', instead of 'problem', since there is nothing intrinsically negative in linguistic diversity, but only social, political, or economic issues associated with its management.[4] This phase refers to the public debate about languages in society in the broad sense of the term. The public debate can encompass different public and private actors including the media, associations, employers, activists, and individual influencers. It can concern a question that is perceived by someone as problematic. For example, speakers of a minority language feel discriminated against and ask for public support for their language; the export industry complains about the lack of good skills in foreign languages; families in a region ask for more opportunities for pupils to learn second languages in schools; or human rights activists raise the issue of the social and linguistic integration of adult migrants. Hence, in the public approach to LPP, the social and political context in which new language policies are introduced is the departure point of the analysis, thereby overcoming one of the main criticisms of critical LPP scholars regarding classic language planning (see Martin-Jones and De Costa Cabral 2018: 72).

---

[4] This does not mean that linguistic diversity has not been or cannot be interpreted this way. For example, linguistic diversity has been interpreted as something inherently problematic and impractical in the early stages of the decolonisation process (see introduction, this volume).

This phase is preliminary to the public policy proper because political and institutional actors have not yet taken a clear position on the issue in question. However, it is important to study if, how, and why a certain issue becomes the subject of public debate, in order to understand the origins and the context of language policies. In liberal democratic countries if an important linguistic issue emerges in society, it is likely that sooner or later some politicians will take an interest in it and try to translate it into votes by promising to deal with it. In authoritarian countries, however, the public debate may be absent or heavily constrained.

In LPP research, the study of this phase benefits from the inputs of various disciplines, including sociolinguistics, politics, law, history, economics, educational study, and philosophy. A broad perspective is indeed required to understand why and how a language question becomes a public issue.

3.2   Agenda-Setting

The term 'agenda' comes from the plural of the Latin word *agendum* (i.e., "something to be acted on"). Agenda-setting refers to the process by which public issues come to the attention of governments as something worth acting on. A *public* problem cannot be defined as such if it is not part of a political agenda. During this phase, public authorities debate and propose solutions to the language issue raised. This is a crucial phase of a language policy. In the public policy approach, as shown in Sect. 2, a language policy is such only if it involves the government (in the broad sense of the term, i.e., not only the central government, but also regional and local governments and public bodies).

The agenda-setting phase is eminently political. The implication of this for LPP research is that studying agenda-setting means examining how and why a language policy question has been framed in a certain way, and why it has been placed on the agenda of questions that merit government's attention. Whose interests have been manifested, overtly and covertly? What concepts and notions have been used, possibly coined in that debate? The study of discourse and ideology, therefore, is indispensable to understand the issues at stake. Unsurprisingly, the analysis of political discourse and of policy texts is the object of substantive research in LPP, especially in discursive approaches to LPP (see Wodak 2005; Barakos and Unger 2016). The study of agenda-setting also benefits from

contributions from political science (Cardinal and Sonntag 2015), as well as media studies and social psychology.

3.3   *Policy Formulation and Adoption*

Policy formulation and adoption are at the heart of language policy. Paradoxically, this phase is one of the least studied in much of contemporary LPP research. In this stage, policymakers clarify the goals and means of a policy, how specific measures will deliver certain effects, for whom, under what conditions, at what material and symbolic costs, and delivering what kind of benefits. Policy formulation usually implies the elaboration of alternatives (including the status quo) to address the public issues at hand, and this is followed by the process of selecting of one of them (including doing nothing). The selection of the preferred course of action is simultaneously influenced by various contextual factors, including the degree of urgency and the type of complexity of the issues that the policy is expected to address. Actors involved in policy design and adoption are typically senior civil servants, senior advisors, and, of course, elected politicians who must make the final decisions.

Policy formulation would require an entire chapter on its own. In this section, I focus on some key aspects in the design of language policy ('policy design'), that is, the development of a programme theory, the selection of language policy instruments, and the design of a system of indicators.

The programme theory. The programme theory of a public policy (or programme) is defined as

> the conception of what must be done to bring about the intended changes. As such it is the foundation on which every program rests […] Whether it is expressed in a detailed program plan and rationale or is only implicit in the program structure and activities, the program theory explains why the program does what it does and provides the rationale for expecting that doing so will achieve the desired outcome. (Rossi et al. 2019: 59–60)

The programme theory is also known by different names, e.g., 'theory of change' or 'logic model'. It is a theory about change because it spells out the expected cause-effect relationships between the policy input and its outcomes, that is, how means relate to goals. As noted in Sect. 2, the content of a policy is made of a selection of *goals* and of the *means* to

achieve them. Recall that *inputs* are all means mobilised for the implementation of a policy, e.g., financial, human, and material resources. An *output* is the direct product of a policy, that is, what is funded and achieved (or realised) through the resources allocated to the policy. *Outcomes* or results are the final effects of a policy on the actors directly concerned. For example, the output of a policy aimed at promoting the usage of a minority language through the media is the number of hours of television broadcasting in the minority language, or the number of printed copies of a magazine. The outcome is measured in terms of speakers and actual language use, i.e., in terms of actual viewers of the TV programme and readers of the magazine. When the policy outcomes achieved correspond to the policy objectives, the policy is effective.

The programme theory should not be understood as a rigid, technocratic, deterministic, top-down policy plan that the decision-maker drops on the sociolinguistic reality from above, as sometimes has been mistakenly argued in critical LPP (see Martin-Jones and De Costa Cabral 2018 for an overview). Rather, the programme theory is a working tool to prevent language policy from becoming a disconnected list of individual measures with undefined effects. Preparing a programme theory is not a mechanical exercise, its development interacts with political imperatives from governments, senior civil servants, and political parties, and it is influenced by path dependencies (Cardinal and Sonntag 2015) and policy legacies (Kraus 2018). It is compatible with 'bottom-up approaches' to public policy, as the experience of policy co-design has shown (Blomkamp 2018).

Assessing the programme theory of a language policy entails explaining how inputs are logically (or at least plausibly) connected with outputs and with outcomes, that is, how the language policy measures are expected to influence the evolution of the practices and attitudes of the target actors. A programme theory is a central part of any public policy because it should spell out the assumptions about how resources and activities will lead to expected results. As Rossi et al. note,

> if the program's goals and objectives do not relate in a reasonable way to the social conditions the program is intended to improve, or the assumptions and expectations embodied in the program's design do not represent a credible approach to bringing about that improvement, there is little prospect that the program will be effective. (Rossi et al. 2019: 60)

While simplistic assumptions about people's linguistic behaviour and representation should be avoided, "no major insight is gained by simply mentioning the fairly obvious fact that reality is complex and that the density of interconnections between processes is boundless" (Grin 2003a: 42).

In policy design, the best way to outline a programme theory is usually by working backwards, that is, starting from the objectives and then designing the path to achieve them given the resources available. The programme theory of a policy can be explicit if a language policy is overt, or implicit in covert language policies. Assessments of the programme theory of language policies can be found in Grin (2003b) for policies aimed at protecting and promoting minority languages, and Gazzola (2016b) for the promotion of multilingualism in the public administration.

Policy instruments. Policy instruments are the concrete means "by which governments attempt to induce individuals and groups to make decisions and take actions compatible with public policies" (Schneider and Ingram 1990: 527, quoted in Landry and Varone 2005: 108). A public policy plan should define a combination of instruments that are suitable to achieve the policy goals. A distinction is usually made between four types of policy instruments, depending on the resource the government uses (Howlett 2019). This follows the taxonomy developed by Christopher Hood (Hood 1986).[5] The first resource is named 'authority'. This resource is used in direct regulations based on prohibitions, obligations, the granting of rights, and the application of sanctions. The second resource the government can mobilise is 'treasury'. There are several examples of policy instruments based on the exploitation of this resource such as grants, subsidies, and taxes. Hood names the third resource 'nodality', and this refers to the fact that the government is a key nodal link in the policy network, and it can receive, disseminate, or hide information in accordance with its position in the informal and formal information channels of the network. There are different examples of policy instruments that use this resource, e.g., setting benchmarking, and organising information campaigns, or, on the other extreme, censorship. The fourth resource consists of the organisational capacity of the public apparatus, which can provide goods and services directly (typically

[5] The taxonomy is generally known by the acronym "NATO" for nodality, authority, treasury, and organisation.

via the public administration) or contract out to entities linked to it by contractual ties.

In Gazzola (2021), I have defined 'language policy instruments' (LPI) as the set of policy instruments used in LPP, that is, the means by which governments deliberately attempt to influence the status, the acquisition, and the corpus of a language with a view to inducing individuals and groups to shift their linguistic practices and attitudes in a certain direction that is consistent with the objectives of a language policy. The direct language regulation of labels and commercial signs is an example of a language policy instrument using the 'authority' resource. Subsidising publications in a minority language is an example of a language policy instrument using treasury. An information campaign to reduce the stigma associated with the use of a minority language—e.g., informing speakers about the possibility of using it for official purposes—is an example of a policy instrument using the nodality resource. Bilingual front offices are a straightforward example of this language policy instrument using the resource 'organisation'. The instruments of language policy, therefore, are by no means limited to legislative directives, and 'command and control'-types of policy instruments.

Usually, any language policy makes use of a complex set of LPI. A policy of promoting bilingualism in the public administration of a country or region, for example, may be employ instruments of a regulatory nature such as giving preference to candidates who have a certificate of knowledge of both official languages (Gaspard 2019). A financial instrument such as the bilingualism bonus (i.e., an annual salary bonus for bilingual staff) is implemented in Canada, Belgium, Slovenia, and some regions in Northern Italy to promote language acquisition among civil servants (Maltais 2018; Mazzacani 2021). Free in-house language training for officials who must learn a co-official language is an organisational tool. An awareness-raising campaign to solicit applications for employment in the civil service from native speakers of a minority language is an example of a language policy instrument that uses an information-persuasive mode. For a discussion about language policy instruments from a comparative perspective, see Cardinal et al. (2015), Gazzola (2021), and Cardinal (2023, forthcoming).

Indicator system. Indicators are a very important component of policy design. Without a system of indicators that translate data into meaningful information, we are not able to check whether the policy is developing

as planned and is leading to the expected results. A system of indicators requires a corresponding data collection system to populate them, e.g., via sociolinguistic surveys (see Dell'Aquila and Iannàccaro 2019 for a discussion of quantitative methods in sociolinguistic data collection). The information system of a language policy is defined as the set of indicators used and the procedures for collecting and processing data. A system of indicators must include input, output, and outcome indicators. The latter are the most common type of indicators because they refer to the objectives of a language policy, e.g., promoting the knowledge and use of a language in society.

A broad language policy usually requires a sufficiently comprehensive set of indicators that reflect and represent the different dimensions of the complex target variables, for example, linguistic vitality and language use. To date, to my knowledge, the most comprehensive language indicator system in the world is the *Sistema d'Indicadors Lingüístics*, SIL, produced by the Catalan regional government in Spain (Bruguera Soler et al. 2022).

The methodology of indicator design consists of breaking down complex variables into a set of simpler dimensions that can eventually be translated into measurable indicators (see Gazzola and Grin 2017: 97–100 for details). The variable 'language use', for example, can, in turn, be broken down into 'number of speakers' and 'frequency of use' and these sub-variables can in turn be articulated into areas such as family, friends, work, and official uses (Iannàccaro and Dell'Aquila 2011). As Maggino and Zumbo note,

> the process of measurement in the social sciences requires a robust conceptual definition, a consistent collection of observations, and a consequent analysis of the relationship between observations and defined concepts. The measurement objective that relates concepts to reality is represented by indicators. From this perspective, an indicator is not a simple crude bit of statistical information but represents a measure organically connected to a conceptual model aimed at knowing different aspects of reality. (2012: 202)

In other words, indicators are not mere numbers; rather, they are measuring tools that acquire meaning within a given programme theory.[6]

---

[6] Indicators are the subject of a vast literature in the social sciences, see Land et al. (2012) for an introduction.

To conclude, the study of the third phase of the language policy cycle benefits from input from any relevant discipline in the social sciences and humanities that can help to examine existing policy plans or to assist decision-makers in the formulation of alternatives. The list includes applied linguistics, law, sociolinguistics, educational sciences, economics, public administration, political sociology, and information and communication technologies (see Leoni, this volume).

3.4   Implementation

The implementation phase refers to how governments put policies into effect. Implementation is the set of processes which, after formulation and adoption, are aimed at achieving the objectives of the language policy. The implementation produces outputs and outcomes. Policy implementation can involve the development of specific action plans to steer and manage the implementation of language policy. Such plans concretely indicate who should do what, when, and how.

It would be misleading, however, to look at this phase as a mere execution of predetermined instructions. Firstly, implementers, in particular middle-level managers such as school principals and heads of units, often benefit from a certain degree of autonomy and flexibility in deciding how to administer policy in a concrete way. Secondly, the successful implementation of a public policy requires the cooperation and the direct involvement of various actors such as simple public servants, officers, teachers, and private actors or non-governmental organisations and associations to which the execution of the policy has been contracted out. These actors ultimately implement a language policy plan on the ground. An excellent programme can fail if due attention is not paid to the implementation procedures, to the training and empowerment of those involved, and to the clarity of communication and information flows between those who lead and those who implement the policy. The resistance of civil servants to policy implementation can hinder the success of a language policy (see Mévellec and Cardinal 2020 for a discussion).

The study of the implementation phase of language policies is often the target of ethnographic studies in the field of LPP (see Johnson, this volume). Considerable attention has been given to the study of the actual practices employed during policy implementation, the ways in which actors interpret and appropriate policy plans, and to the potential tensions

between such plans and their actual realisation, in particular in the classroom and in the workplace (Hornberger and Johnson 2011; McCarty 2015).

The collection of data to monitor and control the implementation of the policy is called 'monitoring'. Data can be employed to populate indicators. The study of this phase includes the examination of the practical conditions of implementation and its operationalisation. It requires that particular attention be paid to matters of organisation and delivery, because language policies are typically implemented through an institutional system centred on public administration, often in interaction with other agents.

3.5   Evaluation

Evaluation refers to the processes by which the results of policies are assessed by governmental or societal actors. A range of disciplines is involved in the process of evaluation in addition to the sociolinguistic analysis of the language practices and representations of the members of policy target groups. Legal studies, for example, focus on the ex-post formal evaluation of compliance of the policy with the legislation in force. Technical policy evaluation focuses on the retrospective analysis of problems in policy design and implementation, and on the empirical estimation of outcomes, benefits, and costs. Both qualitative and quantitative techniques can be useful in this respect. Using data collected during monitoring (and possibly other external sources), technical policy evaluation provides a final judgement on the policy based on some relevant criteria. The most important ones are effectiveness, efficiency, and equity or fairness. Evaluating the effectiveness of a language policy means understanding whether and to what extent the set of objectives have been achieved. In this regard, it is important, if possible, to have reliable data and robust analytical tools and techniques to separate the effects of the policy from other concomitant causes (see Rossi et al. 2019 for an overview). The criterion of efficiency (here interpreted as cost-effectiveness) refers to the relationship between the results obtained and the resources employed. Assessing fairness, in the public policy approach, does not mean carrying out an ethical examination, but rather identifying the distributional consequences that the language policy has had for stakeholders. An ex-post evaluation of the equity of a language policy plan in education, for example, could examine whether teaching a minority

language in public education has been homogeneous or whether it has neglected certain groups; whether all pupils have benefited more or less equally from the lessons; or whether children from disadvantaged social and family backgrounds have had difficulties in accessing schools or online courses.

In certain countries, the periodic technical evaluation of language policies is required by national or regional laws or regulations, see, for example, Art. 8 of the "Swiss Ordinance on the National Languages and Understanding between the Linguistic Communities" (Languages Ordinance, LangO) and Art. 25 and 29 of the Regional Law 18/12/2007 of the Italian Autonomous Region Friuli-Venezia Giulia ("Norms for the protection and the promotion of the Friulian language"). See also the sixth section of the Canadian "Policy on Official Languages".[7]

Evaluations can also be carried out by independent academics as part of their research. Papers in this area deal with a variety of cases, including policies aimed at supporting minority languages (Grin and Vaillancourt 1999; Grin et al. 2003; Wickström 2007, 2021; Strassoldo 2016), the evaluation of the efficiency and fairness of language policy in multilingual states (Pool 1991; Leblanc Desgagné and Vaillancourt 2016) and supranational organisations (Gazzola 2006, 2016a; Ginsburgh and Weber 2011), or the impact of education reforms on the labour market and human capital formation (Cappellari and Di Paolo 2018; Ramachandran 2017). For an overview of the literature, see Gazzola et al. (2016).

4   Language Policy Levels and Public Policy Types

It is worth discussing two questions that better locate the public policy approach to LPP in the relevant academic literature. The first concerns the level of analysis, that is, on the one hand the relationships between public language policy and, on the other hand, other forms of handling multilingualism at the level of organisations. The second question is whether there are some general aspects that make language policy unique in comparison with other types of public policies.

The term language policy is sometimes used to describe the decisions and regulations of private and public non-governmental organisations. Thus, the term 'corporate language policy' describes rules and guidelines

---

[7] See https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=26160.

that are "developed specifically for a business organisation or a unit within an organisation, for example, the language policy of the customer service department of a company" (Sanden 2015: 1100). Language policy in higher education refers to the set of rules, decisions, and course of action adopted to deal with linguistic diversity in a university (see examples presented in Conceição 2020; Vila Moreno and Bretxa 2014). Note that I am not referring here to government language policy regarding higher education institutions, e.g., South Africa's *Language policy framework for public higher education institutions* adopted in 2020 by the Department of Higher Education and Training.[8] I am referring to how single universities deal with linguistic diversity in their various activities, including teaching, administration with students, dissemination of research results, and communication with the public. Language policy in the administration of higher education has become an important topic of research in the last decade, as a result of the process of 'internationalisation' of universities, which in practice has often entailed the replacement of national languages with English in both teaching and research (Conceição and Caruso 2022; Wilkinson and Gabriels 2021; Gazzola 2018; Maraschio and De Martino 2012; Hultgren et al. 2014).

From a theoretical point of view, however, the language policy of corporations and organisations is not very different from other forms of internal regulations, e.g., personnel policies. Personnel policies establish internal rules, procedures, and guidelines for hiring employees, governing employee responsibilities, and dealing with employment problems, such as insubordination or discrimination. This type of policy is not a form of public policy but a corporate or organisational response to a certain need.

In general, there are two aspects that distinguish governments' language policy from the language policy of organisations, i.e., the scope of the policy and its instruments. Firstly, the range of language issues addressed in the government's language policy is much broader than language policies at the level of corporations, non-profit organisations, and research institutions. Official language policy can pursue broad goals that, at least from the point of view of political actors who initiate and drive the policy process, are relevant at societal level (whether national or local). Relevant examples include establishing official bilingualism or multilingualism, promoting and supporting minority languages, teaching

---

[8] See   https://www.gov.za/sites/default/files/gcis_document/202011/43860gon1 160.pdf. See also Ngcobo and Barnes (2021).

foreign or second languages, integrating the language of adult migrants and refugees, and promoting a national language at international level. It is important to recall (see also Sect. 2) that the term 'government' does not solely encompass the central executive. The language policies of businesses and organisations, by contrast, have more focused and limited goals, and they are typically a by-product of decisions aimed at pursuing the core goals of the organisation (see Grin and Gazzola 2013 for a discussion). For example, the business of businesses is business, that is, pursuing profit. The statutory goals of a university (note, I am referring to individual universities not academia as a whole) are to produce and disseminate knowledge, and to attract students. Corporate language policy and the language policy of universities are functional to the achievement of these institutional goals.

Secondly, the state can use a broader set of policy instruments that corporations or organisations simply do not have at their disposal. The most important one is the legitimate use of coercion. Government has the capacity to set laws, orders, and regulations and to enforce them through the police and tribunals. Government can levy taxes. By contrast, the scope and the set of interventions relating to language matters that corporations and organisations can develop are generally more limited, and are also constrained by existing laws and norms.

The relationship between language policy at the macro-level (e.g., the national government's official language policy) and the meso-level (e.g., corporate language policy) is a relevant and important subject in LPP research, and the reason for this lies in the complex interplay between the two levels. As Grin shows (2022), the complexity does not stem from the fact that language policy is 'difficult', but rather from the fact that the incentives and constraints faced by actors at the macro- and meso-levels are not necessarily aligned, and, indeed, they might clearly diverge. For example, some private corporations and universities may have more to gain from linguistic uniformity (e.g., promoting the use of a lingua franca in business and research), while society as a whole has an interest in preserving and promoting the vitality and the use of the national/local languages in all domains of social life.

I do not deny that private actors or individuals can undertake initiatives to plan a language, and that these initiatives can be influential. Lexicographers, for example, are involved in language planning by providing guidance regarding usage and perceived correctness. The creation of international auxiliary languages such as Esperanto or Interlingua is

another example of extreme language planning (see Gobbo, this volume). These cases, however, usually are restricted to corpus planning. Single private citizens alone do not have the means to design and implement decisions that significantly affect the allocation of the functions pertaining to a language at the societal level.

I conclude this section with some remarks about the specific nature of language policy as opposed to other forms of language policy. In policy studies, a distinction is often made between *substantive* and *institutional* policies. A substantive public policy aims at solving a problem politically defined as collective in nature, while an institutional public policy relates to the transformation of government organisation, e.g., a change in the structure of public offices or a re-organisation of human resources (Knoepfel et al. 2007). Institutional public policies aim at creating the conditions for the accomplishment of the tasks of the state, including the implementation of substantive policies. This distinction, however, is not always straightforward in LPP. While teaching foreign languages in schools to improve pupils' skills is a clear example of substantive language policy, the choice of a set of official and working languages and the implementation of multilingualism in the public administration of a bilingual country have *both* substantive and institutional elements.

A second important difference between language policy and other forms of public policy, such as the provision of unemployment benefits, public pensions schemes, or healthcare consists in the fact that there is no zero option in language policy. Although governments may claim it practices benign neglect with regard to religions, it cannot practise it with regard to languages because governments must decide on a language in which they will conduct their business (Kymlicka 1995; De Schutter 2007). In legal and military domains (i.e., the army, police, courts, and prisons), as well as in public administration (i.e., tax office and the register, the language used on banknotes, and names of places), the government exerts an exclusive competence, meaning that private actors are not allowed to provide goods and services unless they are regulated by the government. In other important domains of social life such as education and health care, government often plays a central role. Publicly provided services in these domains require the use of at least one language to be delivered. The choice of which languages to use, however, is not a neutral act in as much as the population of a territory speaks different native languages (May 2005). As a result of the pervasive role of government in modern societies, it is misleading to present LPP as being opposed

to a *laissez-faire* stance on language matters (see Gazzola 2014 for a discussion). The correct distinction is only between different degrees of language policy intervention.

5   Conclusions

This chapter provides an introduction to the study of LPP from a public policy studies perspective. It presents the public policy approach to LPP and its relationship with other approaches in LPP research. It explains how the policy cycle model can be used both to study and to plan language policy. This model is suited to organise knowledge about the practice and the process of language planning, i.e., the design, implementation, and evaluation of language policies.

It is desirable for LPP as a discipline to pay more attention to the theoretical and methodological contributions of policy sciences. Firstly, there is a clear demand for such expertise from public bodies called upon to plan, and implement and evaluate language policies. Secondly, as noted by Thomas Ricento (2005: 11), among others, in order to advocate specific policies or policy directions, scholars and decision-makers ought to be able to demonstrate empirically and conceptually the advantages and disadvantages of language policies as well as their distributive effects on the relevant groups of speakers.

References

Ammon, Ulrich, Dittmar, Norbert, Mattheier, Klaus J., and Trudgill, Peter (eds.) (2004). *Sociolinguistics. An International Handbook of the Science of Language and Society* (Vol. 1). Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110141894.1

ARLeF (2021). *Piano Generale di Politica Linguistica per la Lingua Friulana 2021–2025*. Udine: Agenzia Regionale per la Lingua Friulana (ARLeF)—Regione Autonoma Friuli-Venezia Giulia.

Baldauf, Richard B. (1994). 'Unplanned' language policy and planning. *Annual Review of Applied Linguistics*, 14, pp. 82–89. https://doi.org/10.1017/S0267190500002828

Barakos, Elisabeth and Unger, Johann (eds.) (2016). *Discursive Approaches to Language Policy*. Basingstocke: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-53134-6

Bayley, Robert, Cameron, Richard, and Lucas, Ceil (eds.) (2013). *The Oxford Handbook of Sociolinguistics*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199744084.001.0001

Blomkamp, Emma (2018). The promise of co-design for public policy. In Michael Howlett and Ishani Mukherjee (eds.), *Routledge Handbook of Policy Design*, pp. 59–74. London: Routledge. https://doi.org/10.4324/9781351252928-4

Bruguera Soler, Carla, Ferret Baig, Anton, and Torrijos López, Anna (2022). El Sistema d'Indicadors Lingüístics: diagnòstic i avaluació en política lingüística. *Treballs de sociolingüística catalana,* 32, pp. 247–260.

Caldas, Stephen J. (2012). Language policy in the family. In Bernard Spolsky (ed.), *The Cambridge Handbook of Language Policy*, pp. 351–373. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511979026.022

Canagarajah, Suresh (2005). Ethnographic methods in language policy. In Thomas Ricento (ed.), *An Introduction to Language Policy. Theory and Method*, pp. 153–169. Oxford: Blackwell.

Cappellari, Lorenzo and Di Paolo, Antonio (2018). Bilingual schooling and earnings: Evidence from a language-in-education reform. *Economics of Education Review,* 64, pp. 90–101. https://doi.org/10.1016/j.econedurev.2018.03.007

Cardinal, Linda (2023, forthcoming). Language policy instruments. In Michele Gazzola, François Grin, Linda Cardinal, and Kathleen Heugh (eds.), *The Routledge Handbook of Language Policy and Planning*. London: Routledge.

Cardinal, Linda, Gaspard, Helaina, and Léger, Rémi (2015). The politics of language roadmaps in Canada: Understanding the conservative government's approach to official languages. *Canadian Journal of Political Science/Revue canadienne de science politique,* 48 (3), pp. 577–599. https://doi.org/10.1017/S0008423915000517

Cardinal, Linda and Sonntag, Selma K. (eds.) (2015). *State Traditions and Language Regimes*. Montreal: McGill–Queen's University Press.

Conceição, Manuel Célio (2020). Language policies and internationalization of higher education. *European Journal of Higher Education,* 10 (3), pp. 231–240. https://doi.org/10.1080/21568235.2020.1778500

Conceição, Manuel Célio and Caruso, Elisa (2022). Higher education language policies for mobility and inclusion. In François Grin, László Marácz, and Nike K. Pokorn (eds.), *Advances in Interdisciplinary Language Policy*, pp. 215–234. Amsterdam: John Benjamins.

Cooper, Robert L. (1989). *Language Planning and Social Change*. Cambridge, UK: Cambridge University Press.

Dell'Aquila, Vittorio and Iannàccaro, Gabriele (2019). The geolinguistic treatment of demolinguistic data. In Jeroen Darquennes, Joseph C. Salmons,

and Wim Vandenbussche (eds.), *Language Contact. An International Handbook*, pp. 523–535. Berlin/Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110435351-04

De Schutter, Helder (2007). Language policy and political philosophy: On the emerging linguistic justice debate. *Language Problems & Language Planning,* 31 (1), pp. 1–23. https://doi.org/10.1075/lplp.31.1.02des

García, Ofelia, Flores, Nelson, and Spotti, Massimiliano (eds.) (2017). *The Oxford Handbook of Language and Society*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190212896.001.0001

Gaspard, Helaina (2019). *Canada's Official Languages. Policy Versus Work Practice in the Federal Public Service*. Ottawa: University of Ottawa Press.

Gazzola, Michele (2006). Managing multilingualism in the European Union: Language policy evaluation for the European Parliament. *Language Policy,* 5 (4), pp. 393–417. https://doi.org/10.1007/s10993-006-9032-5

Gazzola, Michele (2014). *The Evaluation of Language Regimes. Theory and Application to Multilingual Patent Organisations*. Amsterdam: John Benjamins. https://doi.org/10.1075/mdm.3

Gazzola, Michele (2016a). Multilingual communication for whom? Language policy and fairness in the European Union. *European Union Politics,* 17 (4), pp. 546–569. https://doi.org/10.1177/1465116516657672

Gazzola, Michele (2016b). Programmazione e controllo della 'politica del plurilinguismo' nell'amministrazione federale Svizzera. *Studi Italiani di Linguistica Teorica e Applicata,* 45 (3), pp. 479–497.

Gazzola, Michele (2018). Les classements des universités et les indicateurs bibliométriques: Quels effets sur le multilinguisme dans l'enseignement et la recherche? In Françoise Le Lièvre, Mathilde Anquetil, Lisbeth Verstraete-Hansen, Christiane Fäcke, and Martine Derivry (eds.), *Langues et cultures dans l'internationalisation de l'enseignement supérieur—Languages and Cultures in the Internationalization of Higher Education*, pp. 132–159. Bern/Berlin: Peter Lang.

Gazzola, Michele (2021). *Language Policy Instruments and the Promotion of Multilingualism in the Federal Public Administration of Canada and Switzerland*, Research report no. 21–3. Belfast: REAL—Research group "Economics, Policy Analysis and Language", Ulster University.

Gazzola, Michele, Grin, François, Cardinal, Linda, and Heugh, Kathleen (eds.) (2023, forthcoming). *The Routledge Handbook of Language Policy and Planning*. London: Routledge.

Gazzola, Michele and Grin, François (2017). Comparative language policy and evaluation: Concepts, indicators and implications for translation policy. In Reine Meylaerts and Gabriel González Nuñez (eds.), *Translation and Public Policy: Interdisciplinary Perspectives and Case Studies*, pp. 83–112. London: Routledge. https://doi.org/10.4324/9781315521770-5

Gazzola, Michele, Grin, François, and Wickström, Bengt-Arne (2016). A concise bibliography of language economics. In Michele Gazzola and Bengt-Arne Wickström (eds.), *The Economics of Language Policy*, pp. 53–92. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9780262034708.003.0003

Ginsburgh, Victor and Weber, Shlomo (2011). *How Many Languages Do We Need? The Economics of Linguistic Diversity*. Princeton: Princeton University Press. https://doi.org/10.23943/princeton/9780691136899.001.0001

Grin, François (2003a). Language planning and economics. *Current Issues in Language Planning,* 4 (1), pp. 1–66. https://doi.org/10.1080/14664200308668048

Grin, François (2003b). *Language Policy Evaluation and the European Charter for Regional or Minority Languages*. Basingstoke: Palgrave Macmillan. https://doi.org/10.1057/9780230502666

Grin, François (2022). Principles of integrated language policy. In François Grin, László Marácz, and Nike K. Pokorn (eds.), *Advances in Interdisciplinary Language Policy*, pp. 23–42. Amsterdam: John Benjamins. https://doi.org/10.1075/wlp.9

Grin, François and Gazzola, Michele (2013). Assessing efficiency and fairness in multilingual communication: theory and application through indicators. In Anne-Claude Berthoud, François Grin, and Georges Lüdi (eds.), *Exploring the Dynamics of Multilingualism: The DYLAN Project*, pp. 365–386. Amsterdam: John Benjamins. https://doi.org/10.1075/mdm.2.17gri

Grin, François, Tom Moring, Durk Gorter, Johan Häggman, Dónall Ó Riagáin, and Miguel Strubell (2003). *Support for Minority Languages in Europe*. Brussels: European Commission.

Grin, François and Vaillancourt, François (1999). *The Cost-Effectiveness Evaluation of Minority Language Policies: Case Studies on Wales, Ireland and the Basque Country,* ECMI Monograph n. 2. Flensburg: European Centre for Minority Issues.

Heywood, Andrew (2019). *Politics (5th)*. Basingstoke: Macmillan.

Hood, Christopher (1986). *The Tools of Government*. Chatham: Chatman House.

Hornberger, Nancy and Cassels Johnson, David (2011). The ethnography of language policy. In Teresa L. McCarty (ed.), *Ethnography and Language Policy*, pp. 273–289.

Howlett, Michael (2019). *Designing Public Policies. Principles and Instruments* (2nd). Abingdon: Routledge. https://doi.org/10.4324/9781315232003

Howlett, Michael, Ramesh, M., and Perl, Anthony (2020). *Studying Public Policy. Principles and Processes (4th)*. Oxford: Oxford University Press.

Hultgren, Anna Kristina, Gregersen, Frans, and Thøgersen, Jacob (eds.) (2014). *English in Nordic Universities. Ideologies and Practices*. Amsterdam: John Benjamins. https://doi.org/10.1075/wlp.5

Iannàccaro, Gabriele and Dell'Aquila, Vittorio (2011). Numeri soggettivi. Spunti sulla vitalità linguistica da inchieste e trattamenti quantitative. In Bruno Moretti, Elena Maria Pandolfi, and Matteo Casoni (eds.) *Vitalità di una lingua minoritaria. Aspetti e proposte metodologiche/Vitality of a Minority Language. Aspects and Methodological Issues*, pp. 151–192. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.

Jann, Werner and Wegrich, Kai (2007). Theories of the Policy Cycle. In Frank Fischer, Gerald J. Miller, and Mara S. Sidney (eds.) *Handbook of Public Policy Analysis: Theory, Politics, and Methods*, pp. 43–62. London: Taylor & Francis. https://doi.org/10.1201/9781420017007.pt2

Johnson, David Cassels (2013). *Language Policy*. Basingstoke: Palgrave Macmillan.

Johnson, David Cassels (2018). Research methods in language policy and planning. In James William Tollefson and Miguel Pérez Milans (eds.), *The Oxford Handbook of Language Policy and Planning*, pp. 51–70. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/978019045 8898.013.2

Kaplan, Robert B. and Richard B. Baldauf (1997). *Language Planning. From Practice to Theory*. Clevedon: Multilingual Matters.

King, Kendall A. and Wright Fogle, Lyn (2017). Family language policy. In Teresa L. McCarty and Stephen May (eds.) *Language Policy and Political Issues in Education. Encyclopedia of Language and Education* (3rd ed.). Cham: Springer. https://doi.org/10.1007/978-3-319-02344-1_25

Knoepfel, Peter, Larrue, Corinne, Varone, Frédéric, and Hill, Michael (2007). *Public Policy Analysis*. Bristol: The Policy Press. https://doi.org/10.46692/9781847422309

Knoepfel, Peter, Larrue, Corinne, Varone, Frédéric, and Savard, Jean-François (2015). *Analyse et pilotage des politiques publiques. France, Suisse, Canada*. Québec: Presses de l'Université du Québec. https://doi.org/10.2307/j.ctt 1f1hcw3

Kraus, Peter A. (2018). From glossophagic hegemony to multilingual pluralism? Re-assessing the politics of linguistic identity in Europe. In Peter A. Kraus and François Grin (eds.), *The Politics of Multilingualism. Europeanisation, Globalisation and Linguistic Governance*, pp. 89–110. Amsterdam: John Benjamins. https://doi.org/10.1075/wlp.6.05kra

Kymlicka, Will (1995). *Multicultural Citizenship: A New Theory of Minority Rights*. Oxford: Oxford University Press. https://doi.org/10.1093/019829 0918.001.0001

Land, Kenneth C., Michalos, Alex C., and Sirgy, M. Joseph (eds.) (2012). *Handbook of Social Indicators and Quality of Life Research*. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-2421-1

Landry, Réjean and Varone, Frédéric (2005). The choice of policy instruments: Confronting the deductive and the interactive approaches. In Pearl Eliadis, Margaret M. Hill, and Michael Howlett (eds.), *Designing Government. From Instruments to Governance*. Montreal-Kingston: McGill-Queen's University Press.

Leblanc Desgagné, Maxime and Vaillancourt, François (2016). Distribution of the benefits and costs of Canada's federal official language policy: Results for 2001. In Michele Gazzola and Bengt-Arne Wickström (eds.), *The Economics of Language Policy*, pp. 141–164. Cambridge, MA: MIT Press.

Maggino, Filomena and Zumbo, Bruno D. (2012). Measuring the quality of life and the construction of social indicators. In Kenneth C. Land, Alex C. Michalos, and M. Joseph Sirgy (eds.), *Handbook of Social Indicators and Quality of Life Research*, pp. 201–238. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-2421-1_10

Maltais, Cathy Florence (2018). Bilingualism in the Canadian public service: The value of bilingualism bonus. University of Liverpool, PhD Thesis. Liverpool.

Maraschio, Nicoletta and De Martino, Domenico (eds.) (2012). *Fuori l'italiano dall'università? Inglese, internazionalizzazione, politica linguistica*. Roma: Laterza.

Martin-Jones, Marilyn and De Costa Cabral, Ildegrada (2018). The critical ethnographic turn in research in language policy and planning. In James William Tollefson and Miguel Pérez Milans (eds.), *The Oxford Handbook of Language Policy and Planning*, pp. 93–112. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190458898.013.3

May, Stephen (2005). Language rights: Moving the debate forward. *Journal of Sociolinguistics,* 9 (3), pp. 319–347. https://doi.org/10.1111/j.1360-6441.2005.00295.x

Mazzacani, Daniele (2021). *A Comparative Evaluation of the Bilingualism Bonus Policy in the Public Administration*, Working Paper of the Research Group "Economics, Policy Analysis and Language"—N. 21–6. Belfast: Ulster University.

McCarty, Teresa L. (2015). Ethnography in language planning and policy research. In Francis M. Hult and David Cassels Johnson (eds.), *Research Methods in Language Policy and Planning: A Practical Guide*, pp. 81–93. Hoboken: Wiley-Blackwell. https://doi.org/10.1002/9781118340349.ch8

Mévellec, Anne and Cardinal, Linda (2020). Les régimes linguistiques à l'épreuve du territoire: le cas de l'Ontario francophone. *Revue internationale des francophonies,* 7, pp. 1–35. https://doi.org/10.35562/rif.1058

Mintrom, Michael and Williams, Claire (2013). Public policy debate and the rise of policy analysis. In Eduardo Jr. Araral, Scott Fritzen, Michael Howlett, M. Ramesh, and Xun Wu (eds.), *The Routledge Handbook of Public Policy*, pp. 3–16. Abingdon: Routledge.

Ngcobo, Mtholeni N. and Barnes, Lawrie A. (2021). English in the South African language-in-education policy on higher education. *World Englishes,* 40, pp. 84–97. https://doi.org/10.1111/weng.12474

Oakes, Leigh (2023, forthcoming). The historical development of language policy and planning. In Michele Gazzola, François Grin, Linda Cardinal, and Kathleen Heugh (eds.), *The Routledge Handbook of Language Policy and Planning*. London: Routledge.

Pool, Jonathan (1991). The official language problem. *American Political Science Review,* 25 (2), pp. 485–514. https://doi.org/10.2307/1963171

Ramachandran, Rajesh (2017). Language use in education and human capital formation: Evidence from the Ethiopian educational reform. *World Development,* 98, pp. 195–213. https://doi.org/10.1016/j.worlddev.2017.04.029

Ricento, Thomas (2005). Language policy: Theory and practice—An introduction. In Thomas Ricento (ed.), *An Introduction to Language Policy. Theory and Method*, pp. 10–23. Oxford: Blackwell.

Rossi, Peter H., Lipsey, Mark W., and Freeman, Howard E. (2019). *Evaluation: A Systematic Approach* (8th ed.). Thousand Oaks, London: Sage.

Rubin, Joan and Jernudd, Björn H. (eds.) (1971). *Can Language Be Planned? Sociolinguistic Theory and Practice for Developing Nations*. Honolulu: The University Press of Hawaii.

Sanden, Guro R. (2015). Corporate language policies—What are they. *Journal of Economics, Business and Management,* 3 (11), pp. 1097–1101. https://doi.org/10.7763/JOEBM.2015.V3.341

Schneider, Anne and Ingram, Helen (1990). Behavioral assumptions of policy tools. *The Journal of Politics,* 52 (2), pp. 510–529. https://doi.org/10.2307/2131904

Spolsky, Bernard (2004). *Language Policy*. Cambridge, UK: Cambridge University Press.

Spolsky, Bernard (2012). What is language policy? In Bernard Spolsky (ed.), *The Cambridge Handbook of Language Policy*, pp. 3–15. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511979026.003

Spolsky, Bernard (2019). A modified and enriched theory of language policy (and management). *Language Policy,* 18, pp. 323–338. https://doi.org/10.1007/s10993-018-9489-z

Strassoldo, Marzio (2016). *Economia delle minoranze linguistiche. La tutela della diversità come valore*. Roma: Carocci.

Tollefson, James William and Pérez Milans, Miguel (2018). Research and practice in language policy and planning. In James William Tollefson and Miguel Pérez Milans (eds.), *The Oxford Handbook of Language Policy and Planning*, pp. 1–32. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190458898.013.1

UNESCO (2021). *Global Action Plan of the International Decade of Indigenous Languages (2022–2032)*, 41 C/INF.14. Paris: UNESCO.

Vila Moreno, F. Xavier and Bretxa, Vanessa (eds.) (2014). *Language Policy in Higher Education. The Case of Medium-Sized Languages*. Bristol: Multilingual Matters. https://doi.org/10.21832/9781783092765

Wickström, Bengt-Arne (2007). Fairness, rights, and language rights: On the fair treatment of linguistic minorities. In Pio Baake and Rainald Borck (eds.), *Public Economics and Public Choice. Contributions in Honor of Charles B. Blankart*, pp. 81–102. Berlin: Springer. https://doi.org/10.1007/978-3-540-72782-8_5

Wickström, Bengt-Arne (2021). Dynamics, costs and the survival of minorities: Optimal language policies for increasing the vitality of minority languages. *Minorités linguistiques et société/Linguistic Minorities and Society,* 17, pp. 221–239. https://doi.org/10.7202/1084706ar

Wilkinson, Robert and Gabriels, René (eds.) (2021). *The Englishization of Higher Education in Europe*. Amsterdam: Amsterdam University Press. https://doi.org/10.1515/9789048553914

Wodak, Ruth (2005). Linguistic analyses in language policies. In Thomas Ricento (ed.), *An Introduction to Language Policy. Theory and Methods*, pp. 170–193. Oxford: Blackwell.

Corpus at the Core: The Epistemology of Language Planning

Federico Gobbo

Abstract This chapter addresses the question of the specific epistemic role of language planning within language policy and planning (LPP) and in the larger field of applied linguistics. Drawing from the traditional distinction between internal and external linguistics, the core of language planning is found in the internal dimension of LPP. The empirical insights obtained by language planning permit avoiding the tragedy of inductivism on one side and the taxonomic fury on the other. In other words, when properly considering empirical data from language planning, the design of the corresponding language policies comes along solid lines so that no gross errors are found. The chapter illustrates the method of levels of abstraction, adapted from the Philosophy of Information, to guide the analysis of case studies, which is exemplified by the case study of Esperanto and Hebrew in comparison, focusing on the epistemological standpoints. The illustrated method shows the importance of inspecting the levels of abstraction, to be done independently from considerations of status, management, or policy.

Keywords Language planning · Internal linguistics · Philosophy of information · Case study analysis · Esperanto · Hebrew

73

1   INTRODUCTION

This chapter aims to identify the contours of language planning in its entanglement with language policy, which underlines the expression language policy and planning (LPP). As the Introduction of this volume illustrates, LPP is a multifaceted endeavour that requires interdisciplinary fertilisation. The immediate consequence is that the theoretical foundations of LPP do not form a coherent canon, but they change through the development of the discipline, from classic language planning research to the ethnographic turn (see Johnson and Ricento 2013; Johnson, this volume). The theoretical foundations of a discipline depend on its epistemological standpoint; in other words, where the limits of human understanding within the discipline itself are to be placed, keeping a scientifically reasonable degree of confidence in its results. Without any doubt, LPP deals with (human) languages, and therefore, our departing point should be languages and their epistemological status.

While the cognitive ability to learn languages is carved into our biology as a species (Tomasello 2009), which languages humans learn depend not from natural causes but rather from the social environment they are placed in. Since the foundation of linguistics as a modern discipline by Ferdinand de Saussure and his *Cours*, published posthumously in 1916, a basic distinction in the discipline is internal vs external linguistics (I refer to the translation by de Saussure, 1959, Chapter V, p. 20):

> My definition of language presupposes the exclusion of everything that is outside its organism [sic] or system – in a word, of everything known as "external linguistics." But external linguistics deals with many important things […] First and foremost come all the points where linguistics borders on ethnology, all the relations that link the history of a language and the history of a race or civilization [sic]. Second come the relations between language and political history. Great historical events like the Roman conquest have an incalculable influence on a host of linguistics fact […] An advanced state of civilization favors the development of special languages (juridical language, scientific terminology, etc.). Here we come to a third point: the relations between language and all sorts of institutions (the Church, the school, etc.).

The main interest of linguistics since its foundation by Ferdinand de Saussure was its internal part: according to this perspective, internal linguistics

considers each language under scrutiny a self-standing object of analysis, and it explains the relations between language elements without referring to external factors. Although Saussure acknowledged that he neglected historical change in languages, presenting languages as abstract, static objects, essentially for pedagogic reasons, nonetheless his influence was so great that the different branches of external linguistics looking at languages from the outside, pointing out how external factors influence the internal development of languages—e.g., society, for sociolinguistics, ethnicity for ethnolinguistics—were left at the periphery of the discipline. As Bourdieu (1982: 8–9) put it:

> The whole fate of modern linguistics was decided, in effect, by the inaugural move whereby Saussure separated "external linguistics" from "internal linguistics", and, reserving the name of linguistics for this second discipline, excluded from it all forms of research that put a language in contact with the ethnology or political history of those who speak it, or the geography of the area where it is spoken, on the grounds that these factors would bring nothing to the knowledge of language taken in itself.

From the point of view of language sciences, LPP is part of applied linguistics, and its field of investigation partly overlaps with sociolinguistics and ethnolinguistics in particular (Johnson 2013). However, LPP holds a specific epistemological view of language that cannot be reduced to the other branches of applied linguistics. In the following, I put forward an epistemologically clear foundation of the concept of language at the core of LPP. The main argument is the following: if we treat LPP as a purely external linguistic branch, we lose empirical data to sustain the effectiveness of the actions over the specific case study under scrutiny. In fact, external activities surrounding language use eventually impact concrete, actual use by its speakers—or signers, in the case of sign languages (on this level of abstraction, their respective epistemologies are the same). In other words, no LPP action leaves the materiality of the languages involved as they are, and the analysis of what happens to the languages gives proof of the degree of its effectiveness. An example is given by Corsican new speakers, who are locals reappropriating their language voluntarily, as adults, passing from linguistic insecurity to linguistic ease, where the Corsican language pragmatically fits in a multilingual repertoire, instead of aspiring to a standard given

by native speakers (Jaffe 2015 on Corsican; Iannàccaro et al. 2018 on linguistic (un)ease).

In order to sustain this main argument, I proceed as follows. Section 2 illustrates the problem, by articulating the terminological chaos surrounding LPP. The proposed foundation relies on the method of levels of abstraction, borrowed and adapted from Philosophy of Information (Floridi 2011), to guide the analysis of case studies. In particular, I argue that the method of levels of abstraction permits us to avoid the Scylla of the tragedy of inductivism on the one hand (Sect. 3) and the Charybdis of the taxonomic fury on the other (Sect. 4). In order to illustrate what proposed concretely, in Sect. 5 I put forward an exemplary case study: the (re)vitalisation of Esperanto and Hebrew in comparison. I will focus only on the epistemological standpoints behind the two processes, as examination in full of the two LPP strategies that led to their success is out of the scope of this volume. On the latter, I argue why—considering its internal linguistic dimension—the more transparent name should be Israeli, as advocated for instance by Zuckermann (2020). Both Esperanto and Israeli are successful stories as they gathered stable communities using them in speech, and in many domains of human life, they can be regarded as success stories of LPP. Also, the comparison illustrates why considering a two-language case study helps to understand analogies and difference more clearly than a one-language only.

2   Some Cats Are Grey in the Dark: The Place of Language in Policy and Planning

The epistemology of language planning is a topic seldom addressed in the literature, especially after the juxtaposition with language policy, which eventually led to the formula of LPP. Moreover, in general, critical views on the consequences for research and application of language planning put epistemology at the margin of the interests of linguists (exceptions being Ricento 2000; Hill 2010; Jernudd 2011; Nevkapil and Sherman 2015). I argue that we need to reflect on the theoretical and epistemological issues of language planning, if we want to avoid the limitations of early language planning research, when sociopolitical variables were a priori excluded in the name of objectivism (see Introduction of this volume). In fact, the history of language planning dates from the 1950s; therefore, scholars have at their disposal plenty of data and case studies to let such a reflection be well grounded. Admittedly, in the early decades of

language planning (1950–1980) in particular, many language planning endeavours turned out to be failures, in very different contexts, such as Europe and Asia (Baldauf et al. 2008). Two examples of language planning failures in Europe are Samnorsk and Limba Sarda Unificada, respectively, at the nation-state and regional levels. Samnorsk, also known as Pan-Norwegian, was a variety proposed to overcome the existence of two written language standards in Norway, in the period between 1915 and 1964 (Jahr 2016); its destiny followed the political climate in the country, and ultimately, Samnorsk revealed to be an instrument to divide Norwegians instead of unifying them. On the other hand, Limba Sarda Unificada, or Unified Sardinian, was an attempt at enforcing a language standard introduced in 2001 in Sardinia (Italy), after the regional bill for the promotion of Sardinian language and culture in 1997. Being accused of inequality towards the varieties of the southern part of the island, it was then abandoned and replaced with a lighter, orthographic standard, Limba Sarda Comuna, or Common Sardinian, already in 2006 (Lai 2018). Nowadays, both Samnorsk and Limba Sarda Unificada are of merely historical interest; however, in this context, the history of failures has a lot to teach language policy experts and language planners in particular.

Ultimately, failure or success (see case study in Sect. 5) in language planning can be determined quite straightforwardly by answering the following question: is the proposed conscious change of the language under scrutiny widely used by the speech community targeted by the language planning itself, or is it rejected as 'artificial' and 'constructed'? From an epistemological point of view, the former question presumes another one: how do we apply language planning successfully and avoid the failures of the past? At least, there seems to be a consensus among scholars, activists, and stakeholders on the importance of this question, which is relevant to all the languages that do not have the luxury of being standardised and ready for written and oral use in all domains, which are the majority of the languages of the world. In fact, according to Ethnologue (24th edition), there are currently 7,139 living languages, and 3,074 are "likely unwritten"; moreover, among the 4,065 languages which have developed a writing system, it is not clear if and how much it is in use by the respective speech communities (Eberhard et al. 2022). The absence or instability of a writing system is therefore a clear indicator of a scarce level of language planning.

This chapter considers two arguments in finding the place of language planning within LPP. As presented in the Introduction, the main argument of this contribution is that the internal aspects of a language should not be overlooked in any LPP strategy, regardless of the languages involved. In fact, even the elite club of languages that have not only a stable writing system but also substantial literary histories and solid institutions are changed in their internal aspects, at least in their domains of use, which impacts their vocabulary. In other words, external actions on the languages in a given context always shape their form, material, and contact: it is advisable to take them into explicit consideration while designing the LPP strategy. The argument in support of the main argument is that we need to clarify the key terms in use to avoid terminological confusion and, in doing so, design better policies for planning languages. In the following part of this section, I start from the supporting argument, while the main argument will be illustrated in the following sections.

LPP comprises both aspects of policy and language planning alike. Although there is general consensus among experts that policy and planning are intertwined, there is a lack of clarity on their boundaries, and therefore, it is difficult to identify their specific domains, which is reflected in the discussions among scholars on defining the contours of LPP in general (from the seminal Cooper [1989] to Baldauf [1994] until various contributions in Ricento [2006]) or language policy specifically (Johnson 2003). In particular, starting from the language planning aspect, Hornberger (2006) proposes an integrated framework that individuates three levels of abstractions for LPP: status, acquisition, and corpus planning. This framework is the starting point of the analysis proposed below. It is important to underline the fact that other starting points are possible; this choice is justified because the three levels of abstraction are neat, clear, and uncontested by most scholars in the field.

The most abstract level is status planning, which concerns the uses of the language in a given context, to maintain or enlarge the domains of use (function), reducing their stigma (in the worst case) or augmenting their prestige (in the best case). Status planning evidently overlaps with language policy in many aspects, as any effort in promoting a language as official eventually pertains to the public sphere, and ultimately the public policy of languages (see Gazzola, this volume, for a re-framing of language policy as part of public policy). Moreover, if we compare the standard model by Hornberger (2006) with the classic standard model of

language planning by Cooper (1989), it becomes evident that the second level of abstraction of the latter, i.e., language acquisition, was included in status planning in the former. The extrapolation of acquisition is due to the recognition of its importance that became more evident the more research in LPP was conducted.

The more research in LPP advanced, the more the role of the language users was revealed to be crucial for the success of any (public) policy on language. For this reason, LPP gives special attention to the two faces of language acquisition, i.e., language learning and teaching: after all, language use is just an abstraction of the practical activity of language users in action. In other words, there is no language use without language users, and the conundrum is in their linguistic choices. Putting the speaker at the centre of the analysis seems to be a smart move, and some experts in the field went along this line. For example, Spolsky (2004, 2009, 2019) put language ideologies and beliefs as a crucial part of language policy, arguing that they are the necessary motivation underpinning language acquisition efforts (see also Schiffman 1996). However, an epistemological problem arises here: unlike language policy documents and practices, ideologies and beliefs are social constructs without direct data at disposal. In other words, ideologies and beliefs are not observable, as they live in the speaker's mind; furthermore, speakers usually apply them merely unconsciously. In fact, for this reason, they can be elicited only indirectly, for instance via qualitative, semi-structured, interviews (Briggs 2008), or via questionnaires (Dollinger 2015). Indirectness is the strategy used to avoid or at least reduce the impact of the well-known observer's paradox, which may hamper the epistemological solidity of the observables. As Labov (1972: 209) famously put it: "the aim of linguistic research in the community must be to found out how people talk when they are not being systematically observed: yet we can only obtain these data by systematic observation".

Unfortunately, this is not the only epistemological problem present in the literature so far; in particular, instability in terminology adds another layer of difficulty. The panorama of the scholarly literature is complicated even more by the term language management, introduced by Jernudd and Neustupný (1987) but popularised by Spolsky (2009). While LPP traditionally focuses on the public sphere, language management is also applied in other spheres, such as multilingual families (e.g., Curdt-Christiansen and Lanza 2018) and workplaces such as multinational corporations and global firms (see at least Mughan 2020; Tange

and Lauring 2009). From a theoretical and epistemological perspective, this goes in the opposite direction of the standard model of LPP by Hornberger (2006), conflating status and acquisition planning again, coming back to the classic model by Cooper (1989), framing again status and corpus as two sides of the language planning coin. Instead, language management focuses on the metalinguistic activities, like the production of discourse about language in use, eventually comprising acquisition planning and part of status planning (see, for example, the passionate justification of Language Management Theory advocated by Nekvapil 2007; however, Spolsky 2009 seems to consider 'language management' as a replacement for 'language planning' as a whole). Within this perspective of language management, language planning is framed as a series of steps—more precisely: noting, selection, adjustment, and implementation—of a problem-solving process. As Neustupný (1994: 50) clearly states: "any act of language planning should start with the consideration of language problems as they appear in discourse, and the planning process should not be considered complete until the removal of the problems is implemented in discourse". Nekvapil (2016) attempts to clarify and sum up what he calls the 'language management approach', describing it as one of the possible approaches in LPP, with particular attention to the contribution by Baldauf (2012).

Language policy, language planning, and language management: this abundance of terms reflects a variety of theories, models, and frameworks, though not contradictory, that make the orientation difficult, regardless of whether interested readers and researchers are students, stakeholders, or scholars. For the purposes of epistemological clarity, it would not a helpful strategy to state that the three terms are 'more or less synonymous', which turns out to be a meaningless statement; therefore, it is important to clarify their analogies and differences, especially if a student or researcher approaches the literature for the first time. Analogously, it does not help to switch from language policy to language management and back—e.g., Spolsky (2019) puts language policy again in the title, in parenthesis, after language management. It is clear that the picture is complex, and some cats are grey in the dark. This contribution is not the venue to discuss all the differences in detail. Here, the aim is merely to highlight epistemic analogies and differences in the approaches and, more importantly, to offer a possible solution to this terminological chaos, giving to language planning a solid, small but firm, place that can be accepted by scholars beyond their different approaches.

We can notice that, mostly, the ongoing conversations in the academe focus on the most abstract levels of abstraction of LPP, i.e., what Hornberger (2006) calls status and acquisition planning. On the contrary, corpus planning, which is the most specific and concrete level of abstraction in the discipline, is just left out. In a more precise wording, Migliorini (1957) called corpus planning, with a nice Italian coin, *glottotecnica*, which can be understood quite literally as language engineering, an expression fallen out of use. Thus, 'glottotechnique' is the collection of acts that aim to consciously shape the form and content of the language under scrutiny. Since the review on the current state of the art of LPP by Baldauf (2005), there was an appeal to distinguish macro-, meso-, and micro-levels in language planning, arguing that corpus planning, or language engineering, should be part of micro-language planning. Although this perspective allows us to insert language engineering in the research agenda, it runs the risk of pushing language planners into the background, as their research results become very specific, connected with the case study only, with few possibilities left for generalisation beyond the concrete context of application—in other words, the tragedy of inductivism, as explained in the next section.

I argue that the epistemology of language planning should put at the core of its specificity what Cooper (1989) and then Hornberger (2006) call corpus planning, which corresponds to Migliorini's glottotechnique. In other words, language planning is—or it should be—the accurate description of its inner mechanics of change when steered by conscious intervention, informing the reception or perception of the language from the outside, whatever named status planning, language policy, or management. Putting language engineering at the forefront immediately clashes with the organicist idea that languages are natural. Under the perspective of organicism, languages are part of what Aristotle's view on biology comes under the name of *bios*, our "way of life" (Lennox 2010). Therefore, they should fall under the Law of Nature (for an account of how historically organicism influences language sciences and linguistics, in particular, see Bonfiglio 2010, chapter 4). I argue that organicism is the philosophical standpoint according to which conscious intervention for languages is marginal in the realm of applied linguistics (I'll come back on this point the discussion around Pei 1962 below). Still now, in the twenty-first century, language change falls mainly in the domain of spontaneity, i.e., caused only by the 'natural' use by speakers, who are not conscious of the changes they are implementing, and the role of

linguists is to register changes, the more objectively as possible (see for instance the handbook edited by Chambers et al. 2001). Although it is undoubtedly true that the speech community of a language determines its change by using the language, it is also true that there are political or cultural leaders, or any other respected persons (such as writers or musicians), whose language use influences others', quite often via traditional or social media. In other words, all speakers are equal, but some are more equal than others, to paraphrase Orwell. Moreover, some users have the power (and the responsibility) to make decisions about the destinies of the language of their respective speech community; think, for example, about language academies, for the languages that have them.

3   BACK TO FUNDAMENTALS: AVOIDING THE TRAGEDY OF INDUCTIVISM

Upon closer examination, languages are nothing but formidable abstractions: languages are no creation as there is no biological species called *sylvestris vulgaris lingua*, such as, for instance, Linnaeus' *felis sylvestris catus*, i.e., the European wildcat. In other words, what we understand as language naming is a collection of mutually intelligible idiolects used by a human group that identifies itself—and it is identified externally—as a speech community. Let's note, in passing, that, not by chance, this "has been a troubled term, caught in a number of methodological, epistemic and political cross-currents" (Rampton 2010: 274). In general, scholars focus on the macro-level—i.e., the change of the formal dimension of language status under scrutiny, which is the core of language policy, as illustrated before. On the other hand, to avoid helicopter views, other scholars dissect case studies of specific languages in concrete settings; in other words, they go to the micro-level. Sociolinguistic empirical data collected from case studies may offer in-depth analysis, but they may lead to risky simplifications to advance the discipline as a whole, unless clear theoretical frameworks are in place. This is the tragedy of inductivism: each case study is unique and special, with no hope for comparison.

However, language planning offers a way to overcome this limitation. If we abstract from the external dimensions of languages, already discussed in the previous section, what remains at the core of language identification is the dimension of distance; in other words, what Kloss (1967) called, in German, 'Abstand'. From an epistemological perspective, the concept of Abstand entails that there are at least *two* languages at

stake; consequently, language identification is not an *immanent* concept (i.e., about the inner characteristics of a definite language) but a *relational* concept, as it presumes the existence of two or more languages in comparison. In the first naïve but highly creative period of language planning pioneers, a *fil rouge* is easily found: case studies never involved language X in isolation, but always in relation to at least with language Y. Examples are Haugen (1962, 1966) and Tauli (1968) for the Scandinavian area, or the above-mentioned Kloss (1967), along with Weinreich (1953) and Fishman (1972), involved in the revitalisation of Yiddish.

If we call language X 'minority', epistemologically, this entails the presence of a majority language Y against which language X is contrasted. In other words, a minority language is so because there is a majority language: consequently, the property of 'minorityness' is not an essential trait of language X but a relational concept involving X and Y as the majority counterpart. In order to be recognised as a minority language, X has to be identified by distance from Y, i.e., through Abstand. In some context, the distance is so high that no confusion is possible: exemplary X/Y pairs are Welsh/English in Wales, Azeri/Russian in Azerbaijan, and Quechua/Spanish in various areas of South America (for a profile of each aforementioned language, see Eberhard et al. 2022). I argue that robust case studies should never put under scrutiny language X in isolation, but always in the context of the life of its language users, which in general implies being in contact with language Y. After all, if we look at the history of humankind as a whole, monolingualism is a rare exception while multilingualism is the norm in everyday life in the vast majority of societies (see Edwards 1994). Thus, overcoming what is called since David Hume's classic treaty *A Treatise of Human Nature* the tragedy of inductivism—where each case study is, like Leibniz' monads, talking only to itself—passes through the introduction of the *meso*-level, which stays between the macro (language policy) and the micro (language scrutinised in their variation, in isolation). The meso-level considers the changes of all languages in a given context both in their inner dimension and in their mutual relations, with an attention apart to the conscious interventions on their respective corpora. For example, in the case of Piedmontese, a minority language in Northwest Italy, the expansion of its vocabulary to fit modern needs passes through the use of French for substantive derivational morphology to mark distance from Italian (Tosco 2012).

The meso-level is now illustrated through the method of the levels of abstraction, borrowed from reflections on the epistemology of information and Artificial Intelligence (Floridi 2011, 2008; van Leeuwen 2014). The aim of such an adaptation to the field of language planning is twofold. First, the method should give a conceptual toolbox for reflecting on the *core* of language planning (in the sense already clarified of Migliorini's 'glottotechnique'). Second, it should explain how its consequences impact language planning as a whole. The central tenet of the Philosophy of Information is that information is a social construction; as language is one of the main vehicles of information, it is clear that language is a biological product of our cognitive ability as a species that eventually becomes a social construction. Linguistic analysis starts with the description of raw data, *de re*, which are linguistic material as samples of language use, collected empirically during fieldwork—for example, revitalisation efforts of a minority language. The collection of empirical data is called Level of Organization (LoO): their analysis, e.g., phonological or lexical, depends on the model of reality offered by the theoretical framework that the linguist uses to form observable variables. They may be completely quantitative, such as language use data monitored by an office for national statistics (e.g., Stokes 2013), or entirely qualitative, such as video recordings of linguistic biographies produced by informants (e.g., Melo-Pfeifer and Chik 2020). For the purposes of this contribution, the degree of precision of the observables is not relevant; on the contrary, it is crucial to acknowledge that, once the variable is set, all the rest can be abstracted. The information is therefore hidden, not lost. After the analysis, the LoO becomes a proper Level of Abstraction (LoA), which is any collection of variables based on the empirical data: "one might think of the input of a LoA as consisting of the system under analysis, comprising a set of *data*; its output is a *model* of the system, comprising *information*" (Floridi 2008: 315). The LoA allows the theory to investigate the system according to a certain model representing the observables, which eventually can be expressed via adaptations of the scheme of system-level-model-structure (SLMS) by Floridi (2008). Figure 1 adapts the original Floridi's model, when the system under scrutiny is linguistic and relational at the same time. In other words, from here, a 'System of languages' means that we are considering at least two language systems and measuring their Abstand, all the rest being abstracted, in particular the dimension of language policy (or management, for some authors presented above).
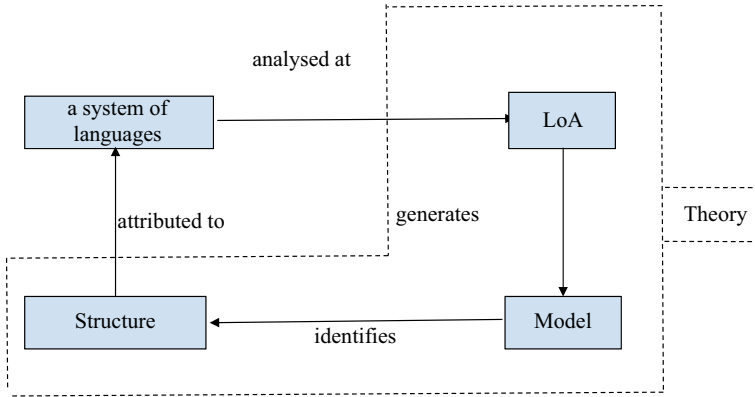
Fig. 1 The method of levels of abstraction applied to language planning

The combination of LoAs generates the model that identifies the structure of the analysed language systems; the whole process is identified as the theory of reference. Figure 1 describes the process of the method of LoAs, not the agents involved, but a LoA is useless if there is nobody to observe the variables. Moreover, observation per se may change the value of the observable variables, according to the position in the quantitative–qualitative continuum. For this reason, the pairs of LoA/LoO are not enough to activate the method just illustrated, and therefore, a third aspect is needed, which is not represented in Fig. 1: the Level of Explanation (LoE), which entails the purposes of the analysis of the case study, and "support(s) an *epistemological* approach, quite common in cognitive and computer science" (Floridi 2008: 319). It is important to note that the LoE comprises all the factors that stay in the eye of the beholder, and for this reason, they cannot be expressed in terms of observable variables. While in computer science LoEs may identify the perspective of end-users, programmers, or customers, in using a determinate software or service, in this context LoEs can identify the language beliefs, attitudes, and ideologies. For example, in applying SLSM to language policy and planning actions in a school, the LoEs will represent covert and overt beliefs, attitudes, and ideologies by the agents involved, such as school pupils, teachers, administration staff, and parents. Continuing the example, the raw data (LoO) may be school documents, video recordings of informal

talks in the playground as well as in the classroom, and so forth. If the case studies involve more languages, as advocated in Sect. 4, the definition of the correspondent LoAs would gracefully take into account the details of the language varieties at stake, including code switching and mixing, if present, identifying the system of languages. In sum, the procedure of performing a language planning analysis should start from the corpus of the languages involved, forming a pile of LoAs, and examining the collected linguistic material per se (LoO) that models the linguistic information identifying the structure of the system of languages under analysis. Structured linguistic information is not a mere collection of disentangled materials anymore, but on the contrary a description of linguistic observables including their mutual relations, that informs the LPP experts (i.e., the LoEs, and therefore out of the language system, as they are not observable variables). It is here that we enter the realm of language policy, which includes large parts of what Hornberger (2006) organises under the rubric of status and acquisition planning.

4   A Procedure to Tame the Taxonomic Fury

Being abstractions, languages invite the multiplication of entities beyond necessity, paraphrasing Occam's razor. I argue that the antidote to this risk is to analyse language systems in comparison. In general, societies are de facto multilingual: therefore, the starting point should be the core of language planning, defined before as corpus planning or glottotechnique. In practice, linguists should calculate the distance (Abstand, following Kloss 1967) of the two or more languages in the system of languages under scrutiny; such a calculation depends on the nature and quantity of the observables at disposal. The goal is to inform language policy experts adequately, before any other action to be taken. In other words, corpus planning should be done before any policy; otherwise, the policy runs the risk of relying on beliefs and ideologies (LoE) instead of observable variables (LoA) rooted in empirical data (LoO), de facto ignoring the reality of the system of languages.

Following Kloss (1967), the notion of Abstand or distance pertains to internal linguistics, as it measures structural distance between two languages (which goes far beyond the lexicon, but involves the structure, such as word order, verbal system, or the presence of the definite article). After delineating the case study of two (or more) languages in the same territory, or context, for instance Frisian (X) and Dutch

(Y) in Leeuwarden (e.g., Gorter et al. 2008), or Aboriginal languages in Australia (Xs) vis-à-vis English (Y; see Zuckermann 2020), it is still unclear how to proceed if Abstand is not properly taken into account. Of course, it is perfectly fine to keep the focus on language X in the minority-majority dynamic; my point here is that, even if in perhaps a small part, the different balancing resulting from an LPP action affects both languages X and Y, as presumably in some domains Y will give space to X, unlike the situation before the LPP action takes place.

In this section, I propose a procedure to perform the epistemic analysis of LPP case studies in five steps (**0–4**) adapted from the case-based approach in computer ethics put forward by Barger (2008). Initially, the linguist should regard the case study prima facie, before to start the analysis. Conventionally, I call it **step 0**. Quite often, the case study was raised to public attention, and a tangle of emotions, feelings, sentiments is inextricably connected to it. In other words, the analyst should identify the degree of his or her involvement and establish the right distance before analysing the case study itself. It is important to collect opinions, attitudes, and ideologies, even if inaccurate, unfair or politically incorrect, by the members of the speech communities as well as by others who are in contact with the languages under scrutiny. The notes taken in step 0 will be reprised while building scenarios (step 3, see below). On a first glance, this may seem to contradict the method of LoAs, as no observables are construed on these raw data. However, on the contrary, this step is often neglected, even by LPP experts, and failing to take prima facie reactions into account can distort the whole analysis (steps 1–4) with an implicit bias. One of the most frequent objections to LPP as a whole comes from descriptive linguists, who are genuinely convinced that linguistics should not enter the realm of judgement, as if they were chemists dealing with test tubes. However, some topics are recurrent and keen to partisanship—paradigmatically, the long-lasting debate on English in a 'glocalised' world (see Gobbo 2015). Already Mario Pei (1962) averted this paradigmatic shift and its dangers:

> The linguist should go beyond setting forth the facts of language as he sees them. Language is a human activity, and therefore subject to intelligent guidance and handling, even more than plagues, iron rails and chemical compounds. Language is primarily a human tool for human use. If blind forces have been allowed to shape that tool in the past, there is no reason why they should be allowed to continue to shape it in the future […] It is

the linguist's task, among others, to shape existing languages so that they may become better, finer, hardier tools for human use, and to try to evolve a form of communication that may eventually lead to world understanding, at first in the purely material sense of the word, later perhaps in that more spiritual, much abused, much misunderstood sense which may ultimately spell out a diminution of conflict, prejudice, hatred, intolerance and war. (Pei 1962: 137–138)

Once emotional reactions and objections to the whole LPP efforts are collected, the analysis can start. **Step 1** is the acquisition of information to identify the current state of affairs. In the scheme of SLMS, step 1 aims to collect the raw data (LoOs) that emerge from fieldwork or existing literature. This seems to be straightforward, but in practice it is not: quite often, the data collected pertain to only one language in the multilingual context and neglect a whole bunch of data, which can potentially change the subsequent analysis. While dealing with the core of language planning, this means collecting real-use linguistic samples of all the languages involved, not only the one under the microscope. Sometimes the absence of correlation in the data is meaningful as well, as the analysis can proceed *ad excludendum*; in other words, the collection of raw data is normally larger than the dataset actually used in the definition of the observables (LoAs). In particular, agency identification is crucial: it is different if linguistic data are uttered by members of the speech communities, traditional or new speakers, policy agents, stakeholders, and so forth. Moreover, raw data comprise not only the brute data produced, but also the process that led to the formation of such data. Of course, it is also possible that some raw data are identified only partially, or that they are not available for the analysis, for a lot of different reasons. If collection does not lead to clear identifiable data, they should be marked as existent but *opaque*; thus, even if not considered, it should be pointed out that they actually exist, for subsequent research and advancement of knowledge in the given case study.

Once raw data are collected and ordered in LoOs, the linguist passes to **step 2**, performing the analysis that forms the "LoAs correspond[ing] to different representations or views" (Floridi 2008: 311), that eventually identify structures—see the scheme SLMS in Fig. 1. If the observable variables are discrete, they can be ordered as a range of values, and they can be called gradients of abstractions (GoAs; Floridi 2008). At the other end of the spectrum, observable variables can be ordered only partially;

mathematically, the models of partial ordering can be formulated in terms of (semi-)lattices (Burris and Sankappanavar 1981). At the extreme, qualitative observables are tag clouds, i.e., a collection of words pertaining the data, without any internal order. Whenever possible, is it better to formulate LoAs that permit at least a partial ordering, as the structure identified by the model results in clearer inspection and scrutiny, as well as comparison with other language systems. **Step 3** looks at the relations between the previously identified LoAs, including agents, and predictable outcomes. Once step 3 is completed, the structure attributed to the language systems is ready for the last step, which concerns evaluation of the model just defined.

Step 4, unlike the previous ones (0–3), is speculative, although grounded in the scheme SLMS just defined. In fact, step 4 is devoted to the formulation of scenarios. The aim is to formulate viable solutions to a new relationship between languages X and Y, evaluating possible outputs of the LPP action. For each LoA, one or more purposes should be identified, according to the agents involved, and their respective commitments. This is the moment in which the notes taken prima facie (step 0) turn out to be useful, as they double-check the correctness and completeness of the analysis done so far. Once all the pieces are put on the board, three scenarios are foreseeable: idealistic, minimalistic, and pragmatist, in this order. The idealistic scenario depicts the best possible solution, when (almost) all agents are satisfied about the new output. On the opposite side, the minimalistic scenario describes an acceptable output that does preserve the most important shared objectives. The third scenario is formulated as a barycentre of the previous ones; in other words, it is a middle way where all language systems should profit at a certain extent and, therefore, all agents involved shall find a reasonable degree of satisfaction.

5   Two Contact Languages: Esperanto and Hebrew (Israeli) in Comparison

This section shows how the levels of explanation (LoEs) illustrate the epistemological standpoints behind LPP efforts. The comparison between the two languages illustrates the importance not to consider each language in isolation. Eventually, such a comparison leads the application of the method of levels of abstraction (LoAs) illustrated via the scheme SLMS above. For reasons of space, the procedure of the epistemic analysis of

LPP case studies will focus on the preliminary step 0. The purpose is to show how the information obtained through 'glottotechnique', i.e., the analysis of the language corpus after successful processes of language engineering, changes the epistemological perspective. This is why at the end I call the language of Israel, instead of the more common 'Modern Hebrew', Israeli.

The case study of the comparison between the LPP process that led respectively to Esperanto and Israeli was chosen under three criteria, which orient our levels of explanation (LoEs, as illustrated above): epistemological, genetic, and scholarly. The three criteria help us to establish step 0 in the procedure proposed in the previous section, i.e., the prima facie reactions. This helps the analyst and the readers understand the positioning of the analyst, so to avoid or at least minimise biases. The first criterion is epistemological. Both cases of Esperanto and Israeli are examples of extreme language planning, especially in its core, i.e., their respective glottotechniques in the corpus. The revitalisation of Hebrew is a spectacular success: ceased to be spoken as an everyday language when Jews adopted Aramaic in the second century CE, it remained a liturgical language until the process of revitalisation in the nineteenth century initiated by Ben Yehuda (Fellman 1973). On the other hand, Esperanto is the most successful planned language ever; launched as a project in 1887, it succeeded in forming a community of practice that naturalised the language, up to the point of having family speakers; for a presentation of the reasons behind its relative success, see Blanke (2009).

Since at least the classic Hagège (1985), linguists from the francophone world, notably Calvet (1998), gave some attention to Esperanto in terms of its language policy scenario—i.e., possible output of LPP strategies where Esperanto actively plays a role, alongside other languages—but they end with positions of realpolitik, arguing that while Esperanto is preferable to English as a global language, the latter became an Esperanto de facto (Hagège 2010). For these authors, the whole analysis was not rooted in empirical data of Esperanto-in-use, but pure speculation is based on beliefs and ideologies. However, it should be admitted that at least Calvet and Hagège, among few other linguists, gave Esperanto a chance, on paper, while the most common attitude is simply to ignore its existence. This attitude may have its roots in the prejudice against LPP as a whole, especially if the organicist paradigm is still in place, which questions the reality of Esperanto as a living language. This argument *ad linguam* happens even if linguists already in the early days of the past

century discussed the phenomenon of planned languages (then still called 'artificial') without any prejudice, most notably the inventor of the fortunate expression 'lingua franca', Hugo Schuchardt (1909). As Pauli (1968: 22) recalls:

> Schuchardt maintained as early as in 1904 that all artificial languages are more or less natural, and the natural languages are more or less artificial. Jespersen reiterated the same in 1929. This means that the difference between a "natural" and "artificial" language is only one of degree, i.e. in all ethnic languages there occur arbitrarily created elements.

On the other hand, the revival of Hebrew, even if it shows striking similarities—as well as evident differences—with the genesis of Esperanto, it is never questioned as a 'real' language, because of the existence of Israel. According to Zuckermann (2020, 2006), the revival of Hebrew eventually brought the genesis of Israeli, which is a contact language that guarantees "continuity not only of liturgical Hebrew, but also of the mother tongue(s) of the founder generation (mostly Yiddish)" (2020: 6). Naming the product of the revival of Hebrew 'Israeli' instead of 'Modern Hebrew' is justified by Zuckermann through a detailed and extended analysis of observable variables that typologically puts the language of Israel not as a direct continuation of the written language in the Hebrew Bible, but rather as a unique—and partially unexpected—output defined by the first generation of first-language users in the early years of Israel (see Zuckermann 2020, 2006 for details). Moreover, the phonology of the language shows considerable traits that corroborate this view as "the most interesting [and] the most accurate", according to Pariente (2021). The internal analysis of the two languages in comparison shows that extensive knowledge of the language situation at stake, including the internal traits, is crucial to express grounded evaluations, especially in the formulation of language policies.

The second criterion for the comparison between Esperanto and Israeli is the genetic level of explanation (LoE): the parallel between the genesis of the two languages. Zuckermann (2020: 164) even advocates that "future research should compare Revival Languages such as Hebrew, and artlangs (artistically-constructed languages) such as Quenya and Klingon". However, the sociolinguistic situation of Esperanto is unique in the panorama of constructed, planned, or 'artificial' languages. While Klingon has a relatively long history of a community of enthusiasts, thanks

to the Star Trek phenomenon, its semi-fluent speakers never went over a hundred, and there is no intergenerational transmission, inside or outside the family (Wahlgren 2021, 2004). On the contrary, Esperanto shows all the traits of a lively and living community for more than a century, whose classification is still a matter of discussion among specialists (Fians 2021; Gobbo 2017; Stria 2015). More importantly, both Esperanto and Israeli show the following traits in their genesis: their respective language architects, Ludwik Lejzer Zamenhof and Ben Yehuda, were Ashkenazim from the Tsarist Empire, involved in the proto-Zionist movement, and well cultivated both in the traditional Hebrew and in the classic and modern European cultures (Tonkin 2015; Lindstedt 2009). Moreover, their native tongues were Yiddish and Russian, which greatly influenced the structural core of both languages. The influence of Yiddish, in particular—despite the effort by the founding father of the language of avoiding it—is the main argument put forward by Zuckermann (2020: 106) to call the language Israeli and not Modern Hebrew: "Israeli is distinct from its contributing parts. To use a chemical metaphor, it can be regarded as a compound or solution, rather than a mixture or suspension". I argue that this chemical metaphor is valid for Esperanto as well.

The genesis of both languages brings us gracefully to the third criterion, which is scholarship in fashion. In particular, research on the comparison of revitalised and planned languages is scant—an exception being Romaine (2011)—while often the comparison is oddly brought towards creoles, even if the two phenomena, i.e., creoles and planned languages, have few traits in common. An example of this struggle is found in Gledhill (2000: 41–42): "Esperanto can be described as an artificial Creole... however, Esperanto has little in common with true Creoles... More realistically, Esperanto can be compared to languages which have been reviewed and extended for social or political reasons, such as Ivrit [another name for Israeli]". While Marcialis (2011) and Lindstedt (2009) rightly point out the importance of the Slavic influence on Esperanto, its typological collocation puts it at the margins of the Standard Average European, but still inside its perimeter. Parkvall (2010: 72) measures the distance of the Esperanto languages system compared to a list of typical traits of European languages and concludes that "Esperanto is indeed more European in character than many of its advocates would have it, but probably less so than many of its opponents would have predicted". Interestingly, it was also argued that (Israeli) Hebrew is a descendant from Yiddish and therefore by the end a Slavic language

(Wexler 1990; see Fellman 1973 for a different view). As announced in the introduction to this chapter, we take the analysis of the structures of Esperanto and Israeli in all its layers (phonetics, phonology, morphology, syntax, lexicon, and phraseology) for granted, as the purpose of this chapter is purely epistemological and the case study only illustrative, not in full. While leaving the discussion of the genealogical and typological collocation of both languages out, we can say that the common ground across the different linguists is that both languages show a high level of hybridisation and therefore should be considered compound languages by contact.

6 Conclusion

In this chapter, I showed how to disentangle language planning from the usual formula of language policy and planning with an argument based on its epistemology, which is based in a need for a solid description of language systems and the conscious intervention on its material at all layers, from phonetics to lexicon and phraseology. The (re)vitalisations respectively of Hebrew and Esperanto are vividly representative, as the they illustrate extreme cases of LPP actions where language planning is at its extreme.

The importance of inspecting the LoAs should be done independently from considerations of status, management, or policy. I argue that the core of language planning, i.e., the rigorous scrutiny of the underlying language engineering methods and techniques, should inform language policy experts, especially if not working linguists, in order to avoid gross errors without any linguistic ground, such as, in this case study "Esperanto is not European because it is a world language", or "the language of Israel is pure Hebrew". Migliorini's (1957) concept of glottotechnique, i.e., language engineering, should be reprised and expanded. By using the method of the levels of abstraction designed within Philosophy of Information, mainly by Floridi (2011, 2008), a solid case-based approach can put forward, adapting main results obtain in the field of Computer Ethics, in particular by Barger (2008).

Case studies should take at least two language systems in comparison, so to measure their distance in terms of Abstand, as defined by Kloss (1967). The aim of such a disentanglement is to correctly inform language policy experts about the reality of the system of languages in use through empirical data and levelled analysis. In this way, analysts

external to the corpus—the core of language planning—whatever they define themselves in terms of language policy, management, or status and acquisition planning, will not run the risk of avoiding unrealistic expectations in the definitions of their actions, in particular in terms of scenario formulation.

References

Baldauf, R. B. and Siew Kheng Catherine Yeo-Chua and Jocelyn Graf and Obaidul Hamid and Minglin Li and Hoa Thi Mai Nguyen and Keiko Ota and Dyah Sunggingwati and Hsin-Fei Wu (2008). *Successes and Failures in Language Planning for European Languages in Asian Nations. 5th Nitobe Symposium*. Tokyo, Japan, August 2007. Tokyo, Japan: European Institute, Sophia University; Japan Esperanto Institute.

Baldauf, R. B. (2005). Language planning and policy research: An overview. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning*. Mahwah, NJ: Erlbaum. 957–970.

Baldauf, R. B. (1994). [Unplanned] language policy and planning. *Annual Review of Applied Linguistics,* 14, 82–89. https://doi.org/10.1017/S0267190500002828

Baldauf, R. B. (2012). Introduction—language planning: Where have we been? Where might we be going? *Revista Brasileira de Linguística Aplicada*, 12 (2), pp. 233–248. https://www.redalyc.org/articulo.oa?id=339829641002

Barger, R. N. (2008). *Computer Ethics: A Case-Based Approach*. Cambridge: Cambridge University Press.

Blanke, D. (2009). Causes of the relative success of Esperanto. *Language Problems and Language Planning*, 33 (3): 251–266. https://doi.org/10.1075/lplp.33.3.04bla

Bonfiglio, T. P. (2010). *Mother tongues and nations: the invention of the native speaker*. New York: De Gruyter Mouton.

Bourdieu, Pierre. (1982). *Ce que parler veut dire. L'economie des echanges linguistiques*. Paris: Fay

Briggs, Charles L. (2008). Sociolinguistic Interviews. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier and Peter Trudgill (eds.), *Volume 2: An International Handbook of the Science of Language and Society*, pp. 1052–1062. Berlin/New York: De Gruyter Mouton. https://doi.org/10.1515/9783110171488.2.7.1052

Burris, Stanley N. and Sankappanavar, H. P. (1981). *A Course in Universal Algebra*. Berlin: Springer.

Calvet, Louis-Jean. (1998). *Language Wars and Linguistic Politics*. Oxford: Oxford University Press.

Chambers, J. K., Trudgill, Peter, and Schilling-Estes, Natalie (eds.). (2001). *The Handbook of Language Variation and Change*. Oxford: Blackwell.

Cooper, R. L. (1989). *Language Planning and Social Change*. Cambridge: Cambridge University Press.

Curdt-Christiansen, X. L. and Lanza, E. (2018). Language management in multilingual families: Efforts, measures and challenges. *Multilingua*, 37 (2), pp. 123–130. https://doi.org/10.1515/multi-2017-0132.

Davey, B. A. and Priestley, H. A. (2002). *Introduction to Lattices and Order*. 2nd ed. Cambridge: Cambridge University Press.

de Saussure, Ferdinand. (1959). *Course in General Linguistics*. Translated by Wade Baskin. New York: Columbia University Press.

Dollinger, Stefan. (2015). *The Written Questionnaire in Social Dialectology: History, Theory, Practice*. Amsterdam: John Benjamins.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.). (2022). *Ethnologue: Languages of the World*. Twenty-fifth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Edwards, J. (1994). *Multilingualism*. London: Routledge.

Fellman, J. (1973). *The Revival of a Classical Tongue. Eliezer Ben Yehuda and the Modern Hebrew Language*. The Hague: Mouton.

Fians, G. (2021). *Esperanto Revolutionaries and Geeks: Language Politics, Digital Media and the Making of an International Community*. Berlin: Springer.

Fishman, J. (1972). *The Sociology of Language; An Interdisciplinary Social Science Approach to Language in Society*. Rowley, MA: Newbury House.

Floridi, L. (2011). *The Philosophy of Information*. Oxford: Oxford University Press.

Floridi, L. (2008). The method of levels of abstraction. *Minds & Machines*, 18, pp. 303–329. https://doi.org/10.1007/s11023-008-9113-7

Gledhill, Christopher. (2000). *The grammar of Esperanto: A corpus-based description*. München: Lincom Europa.

Gobbo, F. (2017). Are planned languages less complex than natural languages? *Language Sciences*, 60, pp. 36–52

Gobbo, F. (2015). Transnational communication in a 'glocalized' world: In search of balanced multilingualism. *Language. Communication, Information*, 10, pp. 196–208. https://hdl.handle.net/11245/1.510554

Gorter, D., van der Meer, C., & Riemersma, A. (2008). Frisian in the Netherlands. *Multilingual Europe: Facts and Policies*, pp. 185–206. Berlin: De Gruyter. https://doi.org/10.1515/9783110208351.3.185

Hagège, C. (2010). Hymne à la diversité des langues. In *Les Cahiers du Musée des Confluences. Revue thématique Sciences et Sociétés du Musée des Confluences*, pp. 45–51. tome 5. La Diversité. doi: https://doi.org/10.3406/mhnly.2010.1504.

Hagège, C. (1985). *L'Homme de paroles: Contribution linguistique aux sciences humaines*. Paris: Fayard.

Hill, L. (2010). Language and status: On the limits of language planning. *Stellenbosch Papers in Linguistics*, 39, pp. 41–58. https://doi.org/10.5774/39-0-3

Hornberger, Nancy H. (2006). Frameworks and models in language policy and planning. In T. Ricento (ed), *An introduction to language policy: Theory and method*. London: Blackwell.

Jahr, Ernst Håkon. (2016). The fate of Samnorsk: A social dialect experiment in language planning. In Michael Clyne (ed.), *Undoing and Redoing Corpus Planning*. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110801729-011.

Iannàccaro, G., Gobbo, F., Dell'Aquila, V. (2018). The assessment of sociolinguistic justice: Parameters and models of analysis. In M. Gazzola, T. Templin, B. A. Wickström (eds.), *Language Policy and Linguistic Justice*. Cham: Springer. https://doi.org/10.1007/978-3-319-75263-1_12.

Jaffe, Alexandra. (2015). Defining the new speaker: Theoretical perspectives and learner trajectories. *International Journal of the Sociology of Language*, 231, pp. 21–44.

Jernudd, B. H. (2011). Language planning from a management perspective: An interpretation of findings. In Ernst H. Jahr (eds.), *Language Conflict and Language Planning*, pp. 133–142. Berlin and New York: De Gruyter Mouton. https://doi.org/10.1515/9783110886580.133.

Jernudd, Björn, and Neustupný, J. V. (1987). Language planning: For whom? In Lorne Laforge (ed.), *Proceedings of the International Colloquium on Language Planning*, pp. 69–84. Québec: Les Presses de L'Université Laval.

Johnson, David C. (2013). Introduction: Ethnography of language policy. *International Journal of the Sociology of Language*, 2013 (219), pp. 1–6. https://doi.org/10.1515/ijsl-2013-0001.

Johnson, David C. (2003). *Language Policy*. Palgrave

Johnson, David C. and Ricento, Thomas (2013). Conceptual and theoretical perspectives in language planning and policy: Situating the ethnography of language policy. *International Journal of the Sociology of Language*, 2013 (219), pp. 7–21. https://doi.org/10.1515/ijsl-2013-0002.

Kloss, H. (1967). "Abstand languages" and "Ausbau languages". *Anthropological Linguistics,* 9 (7), pp. 29–41.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lai, R. (2018). Language planning and language policy in Sardinia. *Language Problems and Language Planning*, 48 (1), pp. 70–88. https://doi.org/10.1075/lplp.00012.lai.

Lennox, James (2010). Bios and explanatory unity in Aristotle's biology. In David Charles (ed.), *Definition in Greek Philosophy*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199564453.003.0010.

Lindstedt, J. (2009). Esperanto—An East European Contact Language? In C. Voß and A. Nagórko (eds.), *Die Europäizität der Slawia oder dies Slawizität Europas. Ein Beitrag der kultur- und sprachrelativistischen Linguistik.* München: Otto Sagner. hdl.handle.net/10138/.

Marcialis, N. (2011). Le fonti slave dell'esperanto. *Studi Slavistici*, VIII, pp. 327–341.

Melo-Pfeifer, S. and Chik, A. (2020). Multimodal linguistic biographies of prospective foreign language teachers in Germany: Reconstructing beliefs about languages and multilingual language learning in initial teacher education. *International Journal of Multilingualism*. https://doi.org/10.1080/14790718.2020.1753748.

Migliorini, B. (1957). *Saggi Linguistici*. Firenze: Le Monnier.

Mughan, T. (2020). Language management in the global firm: Transforming research into education. In Sierk Horn, Philippe Lecomte, and Susanne Tietze (eds.), *Managing Multilingual Workplaces: Methodological, Empirical and Pedagogic Perspectives*, Chapter 10. London: Routledge.

Neustupný, J. V. (1994). Problems of English contact discourse and language planning. In T. Kandiah and J. Kwan-Terry (eds.), *English and Language Planning*, pp. 50–69. Singapore: Academic Press.

Nevkapil, J. and Sherman, T. (2015). An introduction: Language management theory in language policy and planning. *International Journal of Sociology of Language*, 232, pp. 1–12. https://doi.org/10.1515/ijsl-2014-0039

Nekvapil, J. (2016). Language management theory as one approach in language policy and planning. *Current Issues in Language Planning*, 17 (1), pp. 11–22. https://doi.org/10.1080/14664208.2016.1108481.

Nekvapil, J. (2007). From language planning to language management. *Sociolinguistica Jahrbuch (2006)*, 20, pp. 92–104. https://doi.org/10.1515/9783484604841.92.

Pariente, I. (2021). *Theoretical Issues in Modern Hebrew Phonology*. Netherlands Graduate school of Linguistics. University of Amsterdam, Amsterdam Center for Language and Communication.

Parkvall, Mikael. (2010). How European is Esperanto?: A typological study. *Language Problems and Language Planning*, 34 (1), 63–79. https://doi.org/10.1075/lplp.34.1.04par

Pauli, V. (1968). *Introduction to a Theory of Language Planning*, vol. 6 of Studia Philologiæ Scandinavicæ Upsaliensia. Uppsala: Almqvist & Wiksells.

Pei, M. (1962). *Voices of Man: The Meaning and Function of Language*. 1st ed. Routledge. https://doi.org/10.4324/9781003253020.

Rampton, Ben (2010). *Speech Community*. In Jürgen Jaspers, Jef Verschueren, Jan-Ola Östman (eds), *Society and Language Use*. Amsterdam: John Benjamins.

Ricento, T. (2006). *An Introduction to Language Policy: Theory and Method*. Oxford: Blackwell.

Ricento, T. (2000). Historical and theoretical perspectives in language policy and planning. *Journal of Sociolinguistics*, 4, pp. 196–213. https://doi.org/10.1111/1467-9481.00111.

Romaine, S. (2011). Revitalized languages as invented languages. In M. Adams (ed.), *From Elvish to Klingon: Exploring Invented Languages*, pp. 185–225. Oxford: Oxford University Press.

Schuchardt, H. (1909). Die Lingua franca. *Zeitschrift für romanische Philologie*, 33, pp. 441–461.

Spolsky, B. (2019). A modified and enriched theory of language policy (and management). *Language Policy*, 18, pp. 323–338.

Spolsky, B. (2009). *Language Management*. Cambridge: Cambridge University Press.

Spolsky, B. (2004). *Language Policy*. Cambridge: Cambridge University Press.

Stokes, P. (2013). *2011 Census: Detailed Analysis—English Language Proficiency in England and Wales, Main Language and General Health Characteristics*. UK Government. ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/articles/detailedanalysisenglishlanguageproficiencyinenglandandwales/2013-08-30.

Stria, I. (2015). Esperanto Speakers: An Unclassifiable Community? In W. Malec, M. Rusinek, and A. Sadowska (eds.), *Challenging Ideas and Innovative Approaches in Theoretical Linguistics*. Lublin: Wydawnictwo KUL.

Tange, H., and Lauring, J. (2009). Language Management and Social Interaction within the Multilingual Workplace. *Journal of Communication Management*, 13 (3), pp. 218–232.

Tomasello, M. (2009). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Tonkin, H. (2015). Language planning and planned languages: How can planned languages inform language planning? *INDECS*, 13 (2), pp. 193–199.

Tosco, M. (2012). Swinging back the pendulum: French morphology and de-Italianization in Piedmontese. In Martine Vanhove, Thomas Stolz, Aina Urdze and Hitomi Otsuka (eds.), *Morphologies in Contact*, pp. 247–262. Berlin: Akademie Verlag.

Urciuoli, B., and LaDousa, C. (2013). Language management/labor. *Annual Review of Anthropology*, 42 (1), pp. 175–190. https://doi.org/10.1146/annurev-anthro-092412-155524.

van Leeuwen, J. (2014). On Floridi's Method of Levels of Abstraction, *Minds & Machines*, 14, pp. 5–17. https://doi.org/10.1007/s11023-013-9321-7.

Wahlgren, Y. (2021). *The Universal Translator: Everything You Need to Know About 139 Languages That Don't Really Exist*. Stroud: The History Press Ltd.

Wahlgren, Y. (2004). *Klingon as Linguistic Capital: A Sociologic Study of Nineteen Advanced Klingonists*. BA thesis, Soc 346. Supervisor: Bo Isenberg. University of Lund.

Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. New York. https://doi.org/10.1515/9783110802177

Wexler, P. (1990). *The Schizoid Nature of Modern Hebrew: A Slavic language in Search of a Semitic Past*. Wiesbaden: Otto Harrassowitz Verlag.

Zuckermann, G. (2020). *Revivalistics: From the Genesis of Israeli to Language Reclamation in Australia and Beyond*. Oxford: Oxford University Press.

Zuckermann, G. (2006). A new vision for Israeli Hebrew: Theoretical and practical implications of analyzing Israel's main language as a semi-engineered Semito-European hybrid language. *Journal of Modern Jewish Studies*, 5, pp. 57–71.

Schiffman, Harold F. (1996). *Linguistic culture and language policy*. London: Routledge. https://doi.org/10.4324/9780203021569

The Relationship Between Language Policy and Planning, Theoretical Linguistics and Natural Language Processing

Jorge Antonio Leoni de León

Abstract This chapter approaches the convergence of Natural Language Processing (NLP), Theoretical Linguistics (TL), and Language Policy and Planning (LPP). In order to take advantage of resources and opportunities, 30 fundamental questions are given on methodological and ethical matters which will help define the profile of any project and improve the chances of success. The attitudes, expectations, and needs of the language communities should be considered in any technological language project, in order to offer better solutions. In past years, the intertwining between language and technology has increased, and the presence of Information Technology is overwhelming. Therefore, it is important to be conscious that the advantageous technological offer that exists in dominant languages, which also translates into a dominant position, cannot be successfully reproduced in minority or endangered languages. In addition, we must overcome the current situation in which standards are de facto defined by private entities following their own interest. As a consequence, the interest of language communities, especially minority languages and endangered languages, are not necessarily well served. The

development of language applications requires the empowerment of the technology by the linguistic communities themselves.

Keywords Natural language processing · Linguistics · Language applications · Theoretical linguistics · Language and technology · Language sovereignty

1 INTRODUCTION

Language Policy and Planning (LPP) research and practice should pay more attention to the rapid changes in technology and to their potential use in language revitalisation. During the 2020 pandemic caused by SARS-CoV-2, the conditions of isolation imposed accelerated the global digital and technological transformation already underway long before. The need to consider access to web services for all users and, ideally, in most languages has become evident. Natural Language Processing (NLP) is a multidisciplinary area, between linguistics and computer science (in particular Artificial Intelligence), dedicated to the creation of automated tools that allow computers to understand natural language (Charniak 1984; Powers and Turk 1989; Jurafsky and Martin 2008; Manning and Schütze 1999; IBM 2020). For this reason, addressing the relationship between NLP, Theoretical Linguistics (TL), and Language Policy and Planning (LPP) requires us to go through a bit of the linguistic data of the web, given that, also as a result of this transition, we are experiencing the passage from systems based on standalone applications to environments in which online applications predominate. In a general way, we will call them web products.

Indeed, the boundaries between software and web pages are beginning to blur in many cases. For example, software that used to be a standalone application can now be offered as a web application (either on a web page or on a mobile device, such as a mobile phone or tablet, which must be connected to the Internet in order to be used). In this sense, we have concrete examples in word processors that are now offered online with language localization, i.e., linguistically adapted to regions or countries, in a process that can involve a cultural adaptation, for a good number of languages.

Google Docs[1] and Microsoft 365[2] Web applications are the most obvious examples of this evolution. Some web applications for mobile devices can be conceived as extensions of the browser installed by default on the device, but dedicated to a specific web domain (e.g., social networks). Paying more attention to this dimension of technology in LPP is important because, among other reasons that we will detail below, a significant role is played by commercial interests, in which web actors are fundamental, on the vitality and prestige of languages, the perception of them inside and outside the communities, and by the impulse that their use for practical purposes may imply. Software and the Internet have become (and they will increasingly be) important instruments in LPP. It is therefore important to better understand their nature, their role in LPP, and the theoretical and methodological questions that their use arises, in particular in language revitalisation and corpus planning.

This leads us to consider the indisputable reality of the exponential increase of the Internet, and its constant growth, which extends to spheres of public and private life, unsuspected not so long ago, an evolution enhanced by the COVID-19 pandemic. Linguistically, we have experienced the transition from a virtual space dominated, since its origin, by English to a cybernetic universe to which new languages are added every year. Although this change is underway, it does not imply that these languages are used to the same extent and in the same spheres as English. In this process, the dominant languages have benefited from greater representation, so that we have witnessed a progressive integration of the dominant languages (some of them former colonial languages) into the digital networks, long before the languages of their regions of influence. Gradually, other language communities have seen technological alternatives appear in their languages in a process that is often beyond their control.

The Internet has grown steadily since its inception and has evolved from a system for sharing static documents to a complex network of applications and data. In its origins, almost all the documents on the web were written in English, contrary to the current situation where a large number of languages are present. However, although the overall picture gives the impression of a multilingual web, the differences in quality and relevance are remarkable. English still dominates 60% of the web. This percentage

---

[1] https://www.google.com/intl/en/docs/about/.

[2] https://www.microsoft.com.

leaves very little room for the rest of the world's languages (approximately 7000). Most languages on the Internet account for less than 0.8% of documents, services, and applications. The former colonial languages, such as Spanish and French, were integrated into the Internet long before the languages of their respective areas of influence; they constitute what we call the dominant languages and have the highest percentages far below English. There is undoubtedly a multidirectional linguistic inequity on the web.

Percentages can be misleading, because they lead to generalisations. Hence, considering aspects such as the size of the Internet, the relevance and quality of the data, we could come to the conclusion that English, in the first place, and the dominant languages, in the second place, enjoy even greater dominance on the web, far above the other languages. This aspect deserves a great deal of attention, and methodologies will undoubtedly have to be developed for a correct analysis. As far as the subject at hand is concerned, these data give us an idea of the extent of inequalities on the web in linguistic terms.

A truly multilingual web, respectful of linguistic diversity and linguistic rights, requires the creation of resources and products that take into account the grammatical characteristics of each language and the particularities and aspirations of the communities concerned. The very existence of language-based applications is a source of prestige; but, above all, it facilitates the improvement of access to services and information without this being overvalued. This type of information and services are better assimilated in the target community's own language, being its exclusive availability in a foreign or imposed language, a form of exclusion and deterioration of the living conditions of the community.

Computing has evolved in such a way that the boundary between web pages and applications, between data and the products developed from them, has begun to blur, by the very nature of the technologies used today, such as Machine learning, which is based on automated data extraction. Computer applications that were offered as autonomous software in their interaction, now require a constant connection, because they use resources that exceed the capabilities of a simple personal computer. This process of growth has had the effect of including more languages on the web with a consequent adjustment in the standards, defined by private entities, without this always implying an improvement in the quality or relevance of the information in the languages that have been added.

This impulse is largely the result of commercial interests rather than an awareness of language rights and equity in terms of language policy.[3] Let us not lose sight of the fact that companies will prefer to offer products in profitable languages rather than in languages that really require developing their access to technological means. We are talking here about technology not only as a form of entertainment, but also as a means of improving living conditions. We should bear in mind that products and data play a role in imposing standards and procedures, but it is also true that products that are ill-suited to the needs of language communities will hardly be curiosities that will be of much use (see Gazzola, this volume). Still, for many underrepresented languages in computational products (we will call them ULCP from now on) having resources on the web is a form of prestige, but it would be better if it also allows for the creation of products that communities use; in this sense, it is also interesting to mention the relationship between Minority Languages and the Internet, regarding the Oslo Recommendations and the action against digital death (Jackson-Preece 2018).

This chapter revolves around 30 fundamental questions and it is organised as follows. Section 2 takes on linguistics and the language problem, addressing the convergence of NLP, TL, and LPP in the twenty-first century. Section 3 describes the confluence between technology and language, on which 14 questions that every initiative must answer in order to offer products that respond to the needs of linguistic communities are raised. Section 4 addresses the disappearance of limits in technologies and their connection with language and industry, on which 16 questions about fundamental issues that must be answered in any linguistic-technological proposal are advanced. Section 5 concludes. It is important to indicate that some references come from the written press and from non-academic sources and this is due to the frequent lack of support and attention from official media and some academic sectors that affect many linguistic communities around the world.

---

[3] Although there are interest groups on Language Rights and Language Vitality that seek to involve minority languages in current technological development, such as COST Action LITHME (https://lithme.eu), these kinds of efforts are at their beginnings.

2   LINGUISTICS AND THE LANGUAGE
PROBLEM IN THE TWENTY-FIRST CENTURY

The definition of TL is elusive. Determining its contours requires a position that would divert us from our theme, and we do not intend here to decide one way or the other. For us, it is enough to assume the primordial interest of TL with the fundamentals of language. It is important to keep in mind that TL encompasses a large number of sub-disciplines, some more technical than others, but all of them focused on language as a universal phenomenon and its expression in the set of languages of the world. This may give the impression of favouring an increasingly formal approach to language to the detriment of its social aspects. However, TL has been devoted to the description of a great number of languages, many of them disappeared, in its eagerness to apprehend the phenomenon of Language (Hale et al. 1992). TL is not only the repository of an important technical knowledge and documentary heritage, but, as a result of its contact with linguistic communities in search of data, it has also accumulated a great deal of experience in fieldwork, worthy of being taken into account in LPP and NLP.

From a purely technical point of view, without going into deep details about the properties of language and its different manifestations, let us recall that for many languages TL research has made it possible to propose writing systems for several languages. For example, in the Mayan languages (Academia de las Lenguas Mayas de Guatemala, n.d.) and for Costa Rican indigenous languages, such as Bribri (Constenla Umaña 1992; Wilson 1982). The Academy of Mayan Languages is an official entity of the Government of Guatemala. While it is true that the use of such systems has had varying degrees of success, each is the fruit of rigorous work that has required a thorough knowledge of the properties of a language and the creation of a bond of trust with members of the linguistic communities studied. The knowledge produced with respect to a language, in its purely formal interaction or in its everyday use, is undoubtedly valuable for a discipline that must follow a similar path. Hence, these experiences are extremely useful for the development of NLP proposals and for LPP, because the study of the characteristics of a language brings us into contact with linguistic communities that we will get to know better and understand more intimately their expectations.

Writing systems are a basic communication method for NLP, so much so that even speech-based products require some level of written representation.[4] The way in which language communities interact with their writing systems is highly variable and presents particular challenges. A writing system has a form and rules that need to be known and possible to operate with (Bird 2018). A given language community may have one writing system and not use it, or it may have several competing writing systems (recall the case of Bosnian-Croatian-Montenegrin-Serbian (BCMS) where there are two alphabets in use, one Latin and one Cyrillic). It may also happen that the community uses an alphabet of a nearby language and that the orthography is not fixed, as in the case of Swiss German—or Schwyzerdütsch—which uses the alphabet of Hochdeutsch, but with several competing orthographies.

We would like to be able to provide some kind of protocol or roadmap for how to act in cases as those described before, but that is not possible. However, we do want to point out the importance of every linguistic community being involved in the practical processes leading to the representation of their language, such as writing systems, which have a fundamental component in NLP and which also result in Language Planning proposals and Language Policy claims. After all, a writing system is the most identifiable visible element of a language. If we use a writing system in a computational product and it is not used by the linguistic community, the resulting application will be of little value, so understanding the role of such a system in linguistic life is necessary. On the other hand, we can design a Language Planning strategy, but we should not do so without involving the community. As with the issues discussed above, the temptation to copy experiences, attitudes, and methods, particularly those based on the most influential languages, can result in proposals of no benefit. In this sense, an approach to TL knowledge and experience can improve any proposal, particularly if it is sociolinguistically grounded.

---

[4] Originally, computer systems were based on the English alphabet. This was ASCII, an insufficient representation even for languages that use the Latin alphabet such as French and Spanish. The needs of languages other than English prompted the creation of Unicode (https://home.unicode.org), the current universal encoding standard.

3   TECHNOLOGY AND LANGUAGE

The 2020 pandemic caused by SARS-CoV-2 acted as an accelerator of technological change. The isolation conditions boosted the digital and technological transformation. This aspect, which also needs to be taken into account, often involves a linguistic adaptation of computational products known as language localization (Esselink 2000). This adaptation is optional and is a way of adjusting products to make them more easily consumable in different regions. In digital terms, language localization aims, then, to better guarantee the distribution and use of computational products. In a lockdown situation like the one we lived in, that is to say in an isolation imposed as a first response to the pandemic, this process of adaptation was necessary to ensure the quality of services and communication; in particular, if we take into account that this technological transformation implied, in addition to a modification in the use of digital tools and resources, a change in social relations.

From one day to the next, video conferencing, among other applications, became the space for social gatherings, birthdays, parties, and even funerals. It is not uncommon nowadays to have a doctor's appointment or attend a class via video conference. Instant messaging applications came to compete with dynamic websites for buying and selling products. Payment methods were streamlined and the use of cash decreased to such an extent that it practically ceased to be necessary in many locations. The main aspects of this transformation that we can mention are: change in customer needs or expectations (this has consequences in the commercial relationships of each community), increase in remote working or collaboration, raise in customer demand for online purchasing and services, upsurge in use of advanced technologies in operations and in business decision making (LaBerge et al. 2020). While it is true that these concepts are mainly oriented towards commerce, we should extend them to areas such as public administration, education and health, to name a few. Undoubtedly, civil society organisations such as NGOs also faced similar challenges in terms of their daily tasks.

As stated before, NLP is a multidisciplinary area dedicated to the creation of automated tools that allow computers to understand natural language. Unlike Computational Linguistics (CL), which is more interested in modelling natural language, NLP is more concerned in creating resources and tools that would facilitate the understanding of natural

languages by computational systems, independent of linguistic models for the human mind.

The biggest difference between CL and NLP is reflected in the theoretical vocation of CL and in the empirical methodologies of NLP (Llisterri 2003). CL aspires to model human linguistic knowledge and answer theoretically the computational problems it poses. NLP builds linguistic applications, often statistically based. All of these applications are built with methods unrelated to the way humans process language. NLP seeks to solve concrete problems raised by CL (Periñán Pascual 2012), which take shape, for example, in data mining, speech recognition, and machine translation. That said, the two terms are often used interchangeably (Pustejovsky and Stubbs 2013).

NLP has numerous applications in Academia, Science, Civil Society, Government, and Industry. The latter, for commercial reasons, seeks to make its investments profitable in the shortest possible time, hence the generation of resources and methods occupy a central position.

Industry concerns can range from a simple translation of an application to a complete tool based on some degree of understanding of natural languages by computers. The growth of the web has had the effect of making web players aware of the desirability of offering language and regionally adapted products. In some ways, standards are set on the fly by private companies with global interests. It is not always possible to know which criteria prevail in the selection of languages for which they adapt their products.

Perhaps the most notorious NLP technology is Machine Translation (MT). MT is a very competitive area in NLP. Google Translate[5] is undoubtedly one of the best-known tools worldwide (Wu et al. 2016; Le and Schuster 2016). Google Translate allows you to translate texts between a large number of languages, which are selected by peers. The quality of the results varies according to several aspects; one of them is the selection of language pairs: a sentence will be better translated from French to English than its equivalent from Latin to Chinese. Google Translate is an online application, that is to say, it is not possible to instal it as a standalone software on a personal computer, and to use it you need to be connected to the internet. The reason for this is the amount of resources involved in a simple query, which is processed by data centres

---

[5] http://translate.google.com.

(server farms) that coordinate their resources to provide answers to users in the shortest time possible.

The resources needed in NLP, according to the current state of the art, include the compilation of corpora, of the largest possible size (preferably in the order of tens of millions of words). A corpus is a collection of texts compiled for a specific linguistic purpose, whose content may be tagged in a very detailed way (e.g., in its morphosyntactic or semantic features), or it may consist of a series of elementary associated information such as date of compilation, subject, size, etc. This will depend, of course, on the purpose of the corpus. An error in the selection of the corpus or a corpus of insufficient size will have serious consequences on the results of its processing and on the suitability of the resulting product for the purposes for which it was created.

In the area of medicine, for example, it is best to take into account the specialty in which an application will be used, so that the possibilities can be restricted to a very specific domain. Corpora are specific to each language. There are corpora with the same content in several languages; these corpora are called parallel. In the case of Google Translate, parallel corpora make it possible to establish links between paragraphs and sentences that are used to propose translations. From Bible translations, we can build a parallel corpus in a large number of languages,[6] but it would be of little use for the translation of everyday texts, such as news, or specialised texts, such as engineering.

In theory, the methods and algorithms applied to the processing of corpora are universal. That is, they are not language-specific. In practice, it has been shown that it is necessary to adjust methodologies (Casasola Murillo et al. 2017; Casasola Murillo and Leoni de León 2016), because the characteristics of a language may require greater adjustments than those used in the language in which the method was originally developed. For example, a language with a great morphological richness (such as Spanish) requires more particular attention for the treatment of its morphosyntax than a language with poor morphological characteristics (such as English).

These remarks come on purpose, because in many occasions, the development processes of NLP applications for ULCP apply methods

---

[6] A biblical corpus is also interesting, because it allows one to compare ancient versions. One example is the corpus of the Mediaeval Bible Project (http://www.bibliamedieval. es/BM/).

developed, in first instance, for English, without any kind of filter, both for the methods and algorithms, as well as for the compilation of corpora. The first impulse of many research and development teams is to use methods developed for English, applied to any other language, which can be a ULCP, for which results are not obtained with the same quality as for English. It is interesting to note the copying of procedures with a clear lack of attention to the characteristics of the target language. We will not delve into the consequences that this has in the development of NLP applications, but it is important to take it into account, because the same attitude and the same procedure are repeated when collecting corpora.

ULCP do not always have the resources to produce linguistically sensitive applications. We also know that their needs do not always match those of the dominant languages. The expectations of language communities can also be very different. In fact, this varies greatly from one community to another. For example, in Costa Rican languages, a community of a language in the process of obsolescence, such as Boruca or Térraba (brörán t'rocuó), does not have the same needs or the same objectives as one of a language in resistance, such as Limón Creole English or Ngäbere, both from the point of view of educational and cultural needs, as well as the community's knowledge of their language; moreover, access to basic services offered by the government requires a level of proficiency in Spanish in Costa Rican ULCP (Sánchez Avendaño 2013).

It is not uncommon that, on private initiative, as an intellectual exercise, some actor outside a language community produces an application in the language of the community, sometimes even mediated by the language of influence. For example, it could happen that instead of a monolingual online dictionary, a bilingual dictionary is proposed, producing the effect of resorting to a second language as an intermediary to access knowledge in the target language. The problem arises if this bilingual dictionary is aimed at a linguistic community without options in its own language; but it is different if it is proposed along the same lines as any bilingual dictionary, that is, as a way of accessing the lexicon from another language, which means that its audience is different.

Another situation corresponds to the development of an application that will ultimately be only a curiosity. As a result, it will have little or no effect on the language community, because they will make little or no use of the product. Among the causes are that the language has not been used correctly or that the proposed product does not meet any specific need. In this case, the benefit is marginal; the language used or the target

language will not be strengthened from it. It is important to point out that outside the communities, there may be an impact on the prestige of the language by having computational products and resources. However, it is not simply a matter of translating web products, but of adapting them to the needs, expectations, and possibilities of potential users. Such products should even evolve according to the use and transformations in the communities. In short, the methodology and algorithms require technical adjustments with respect to the target languages. But this is not the case with corpora, which are the point of contact with the reality of the languages.

The underlying problem raised by the questions (1)–(14) is related to the frequent separation between the proposal of a linguistic application and the needs of the community to which it is addressed. In general, vulnerable language communities have limited resources to produce computational solutions in their language. In this sense, answering these questions is important in the revitalisation processes in order to optimise the use of resources. Even though there are languages with few speakers that are not threatened, there are dialects or forms of linguistic expressions within the communities that are threatened.

Let us remember that languages are not homogeneous entities, but are made up of a great variety of expressions of which some are preponderant. This is valid for any linguistic community. A priori, it is not possible to rule out that a language considered out of danger does not have a dialect or some form of expression that is in danger. We are not aware of LPP research addressing these questions, but we know that NLP initiatives for endangered languages often ignore the questions we pose, and this causes negative consequences for the projects and, sometimes, even for the language communities, who see valuable resources wasted. In this sense, we are calling the students, researchers, and decision makers to take into account the questions we offer for consideration. The interrelationship between NLP and LPP is, above all, of an ethical nature and this is what the questions (1)–(14) point to:

(1) What is the nature of corpora?
(2) What is the purpose of the collection?
(3) Under what conditions is corpus compiled?
(4) Are community members aware of the use of their data?
(5) What are the characteristics of the compiled corpora?
(6) Is it advisable to anonymize the data?

(7) Is the data really anonymous?

(8) Is anonymity guaranteed?

(9) What are the rights of the language community over the data collected?

(10) Is the corpus for commercial purposes? If it is for commercial purposes, does the community receive any compensation for providing their data?

(11) Who owns the rights to the data?

(12) What is the purpose of the applications developed from these corpora, and to what extent do they benefit the linguistic community?

(13) How does the language community participate in the collection process?

(14) What is the participation of the linguistic community in the products derived from the exploitation of corpora and what are their rights and benefits with respect to them?

As we have already pointed out, the process of corpus compilation is a fundamental stage in the research and development of NLP applications. Corpus collection is not a trivial activity, but an indispensable stage in the development of applications, which must be adjusted, as far as possible, to the linguistic reality of each language. In the case of under-resourced languages, several techniques have appeared in recent years to document them and develop NLP projects (Bird et al. 2014; Blokland et al. 2015; Goldhahn et al. 2016; Stüker et al. 2016; Cunliffe et al. 2021). These are languages for which written documentation is insufficient, as, for example, is the case for the Bribri language of Costa Rica, which has few written texts.

Linguistic reality is better expressed through corpora than through ad hoc examples, as was the custom for many years. Now, the intensive use of corpora should not make us forget their nature and aims, as questions (1) and (2) point out. The purpose of a corpus may be to extract data related to specific domains, such as sport, law, or health. For this reason, the purpose and nature of corpora should be clear to those deciding to work with a given language. The process of building a corpus cannot, in principle, obviate the consent of the contributors, who should be aware of the collection process. Therefore (3), (4), and (5) aim to account for these aspects, which are not always possible to answer in an ideal way, in particular if one uses corpora collected by third parties, in which case

it will be necessary to obtain as much information as possible about the origin of the data and the consent of the participants.

Private data requires particular care, especially if it is health-related information, in which case it must be anonymised. For this, there are standards and tools (Li et al. 2011; Rumbold and Pierscionek 2018; Chakravarthi et al. 2018). However, it always remains open to deanonymise a text (Nanavati et al. 2011; Gambs et al. 2014). If documents are geolocated, they are particularly vulnerable. Hence, questions (6) and (7) address this aspect, while question (8) requires, in addition, a specification of protocols so that third parties cannot deanonymise them or cannot gain access to the most sensitive information. This is extra work, but necessary.

Language is a social phenomenon, hence even though data collection is done with individuals, it has an effect on a whole community. Language is both individual and shared knowledge. Question (9) takes this quality into account; the answer varies dramatically from one language community to another and from one country to another. A small language community may want to retain some rights to the data, while others (the larger ones) do not consider this important. If a community participates in data collection, it can sometimes do so by nominating participants and will expect a concrete return. This may be a recognition of the collective effort, but there may also be other options that need to be specified, particularly if, as question (10) notes, the resulting products are for commercial purposes and there is an issue of compensation, the nature of which may not be financial. Consequently, questions (11) and (12) should have an unambiguous and clear answer for the participants and the consulted community. The answers to the questions in (13) and (14) require that the above questions have been considered by the interested parties. We do not, by any means, claim that the answers are the same for every situation, but it is very important to state them for deliberation.

We consider that work in NLP with ULCP should be stimulated. There are many remarkable initiatives (Bird et al. 2014; Bird 2018; Camacho Caballero and Zevallos Salazar 2020) and the research and development process itself leads us to postulate a scheme that considers the interests and sensitivities of all parties involved. After all, minority language speakers must be allowed access to create content and communicate by digital means in their own language as a basic right (Jackson-Preece 2018). This requires a strong involvement of the community itself based

on its needs, and not be the result of the unilateral action of some agent outside of it.

A framework in which national or transnational companies set the de facto standards is not desirable, nor is it desirable to act as if all language communities had the same visions, expectations, needs, and capacity to act. A vision is needed that, without being relativistic, establishes the responsibilities and commitments of the participants, by which we mean language communities, developers, companies, the government, and civil society. Depending on the project, one or more of these last three actors may be absent.

4   Language Use and Web Products

The current linguistic situation on the Internet, without being ideal, corresponds to that of a multilingual global network, whose percentages reveal an inequitable reality: 62.9% of the web pages are written in English; 7.3% in Russian; 3.8% in Turkish; 3.7% in Spanish; Persian, French, and German share 3.7%, 2.5%, and 2.0% respectively. The remaining percentage is divided by a large group of languages with a representation of less than 0.8% each (W3Techs 2021). Consequently, the amount of information they convey is actually small, and the number of languages represented on the web is low, especially if we take into account the total number of languages in the world, which is about 7139 according to Ethnologue (Eberhard et al. 2021), a figure similar to the data from the University of Maryland's Langscape project (Maryland Language Science Center 2018). If we check the numbers for videos on the YouTube platform, the situation changes a bit. There, English is used in 66.0% of the channels; Spanish, in 15.0%; Portuguese, Hindi, and Korean account for 7.0%, 5.0%, and 2% respectively. The remaining 5.0% is distributed among the other languages (Yang, n.d.). It is important to remark that a large number of people with low educational level prefer to consult videos online (for example, YouTube) instead of researching written sources, such as web pages, of any kind. With respect to English, the change is significant, particularly if we take into account that, although it continues to be the first language of use on the Internet, it no longer maintains 80% of the web content of the 90s (Young, n.d.). We could approach these data in various ways, but for our purposes, we will focus on contrasting them by number of speakers and Internet penetration by language (Table 1 from Internet World Stats).

Table 1 Internet World Stats

Top ten languages used in the Web—31 March 2020 (Number of internet users by language)

Top ten languages in the internet	World population for this language (2021 estimate)	Internet users by language	Internet penetration (% Pop.)	Internet users growth (2000–2021)	Internet users % of world (participation)
English	1,531,179,460	1,186,451,052	77.5%	742.9%	25.9%
Chinese	1,477,137,209	888,453,068	60.1%	2,650.4%	19.4%
Spanish	516,655,099	363,684,593	70.4%	1,511.0%	7.9%
Arabic	447,572,891	237,418,349	53.0%	9,348.0%	5.2%
Portuguese	290,939,425	171,750,818	59.0%	2,167.0%	3.7%
Indonesian/Malaysian	306,327,093	198,029,815	64.6%	3,356.0%	4.3%
French	431,503,032	151,733,611	35.2%	1,164.6%	3.3%
Japanese	126,476,461	118,626,672	93.8%	152.0%	2.6%
Russian	145,934,462	116,353,942	79.7%	3,653.4%	2.5%
German	98,654,451	92,525,427	93.8%	236.2%	2.0%
Top 10 languages	5,273,725,132	3,525,027,347	66.8%	1,188.2%	76.9%
Rest of the languages	2,522,890,578	1,060,551,371	42.0%	1,114.1%	23.1%
World total	7,796,615,710	4,585,578,718	58.8%	1,170.3%	100.0%

Notes (1) Top Ten Languages Internet Stats were updated on January 31, 2020. (2) Internet Penetration is the ratio between the sum of Internet users speaking a language and the total population estimate that speaks that specific language. (3) The most recent Internet usage information comes from data published by Nielsen Online, International Telecommunications Union, GfK, and other reliable sources. (4) Population estimates are based mainly on figures from the United Nations Population Division and local official sources. (5) For definitions, methodology and navigation help, please see the Site Surfing Guide. (6) These statistics may be cited, stating the source and establishing an active link back to Internet World Stats. Copyright © 2020, Miniwatts Marketing Group. All rights reserved worldwide

From Internet World Stats, *Internet world users by language: Top 10 Languages* (https://www.internetworldstats.com/stats7.htm) [Last visit: October 13, 2021]

Note that figures in Table 1 are computed on the basis of a "zero-sum" approach. While many people are bilingual or multilingual, the source used in Table 1 assigns only one language per person in order to have all the language totals add up to the total world population. As Table 1 shows, the main languages of the web, which generally correspond to the colonial languages, once again occupy the first places in number of users, penetration, growth and percentage of users according to world participation, in a way that does not correspond to the population data of their native speakers. For example, English is the mother tongue or the second language (to very different levels of proficiency) of a percentage between roughly 13 to 19% of the world's population, that is, between 1000 and 1500 billion speakers (for example, see Crystal 2006), although we know that this language accounts for 62.9% of the content of the web pages and 25.9% of the users of the network, a fact that will be relevant later on. These data are too general, so we must find the best way to interpret them in order to take advantage of them. This is a much more complicated task and requires us to resort to a series of criteria to clarify the picture.

For this, we need to be familiar with the nature of the technologies that populate the Internet today, the use that linguistic communities make of them, and the linguistic factors that affect their use (consumption) and their creation, without the number of speakers necessarily being the most relevant factor. Within the linguistic factors, we find terminological differences between the dominant languages and the disadvantaged languages; it is not always possible to translate a textbook or even a song; on other occasions, it is not even advisable to translate a commercial text into a minority language, since, after all, the terms to be used are those of the dominant language or it could be that the translated version is not well understood by an important sector of the target linguistic community. Here, translatability and convenience are two key elements that vary from one community to another: while some communities of speakers want to see their language used in all spheres of life, others restrict its use and do not seek to expand or adapt their vocabulary.

However, the data we have presented on languages and the web overgeneralise a reality that is complex. As stated before for questions (1)–(14), we should ask ourselves a series of guiding questions that will help us to find answers about the intertwining of NLP, TL, and LPP. While it is true that there is an imbalance in the relationship between the availability of web products by language and the users, considered according to their

mother tongue, the way in which the data are frequently cited hides other facets of the problem. Thus, we can begin by asking the following series of questions, some of which, although obvious, are necessary:

(15) Does English maintain its dominant position or has it slipped back?
(16) Has the growth of the web allowed underrepresented languages to occupy new spaces?
(17) Does every web product in English have to have an equivalent in other languages?
(18) Does every web product in another language have to have an English equivalent?
(19) Is all relevant information in English relevant in another language?
(20) Is all relevant information in one language relevant in all other languages?
(21) Do web products offered in multiple languages have the same quality and relevance?
(22) What is the quantity and quality of information conveyed by languages that are underrepresented on the multilingual web?
(23) Are web products used in the same way or for the same purposes?
(24) Is the information relevant enough?
(25) Do the numbers relating to the use of dominant languages on the web include non-native users of those languages?
(26) What percentage of underrepresented language users use a language other than their own to create and consume web products, and with what level of understanding?
(27) What percentage of speakers of underrepresented languages are able to use their own language to access web products?
(28) Do web products in underrepresented languages meet the needs of their communities?
(29) Should every language have information on the web?
(30) Should all languages have the same web products?

In the questions (15)–(30), we have used the term "underrepresented languages on the web", it also implies the concept of ULCP. This concept covers created or consumed products, regardless of their platform autonomy, i.e., whether they are standalone or online software, and refers to the set of languages with no or little presence on the Web. Let us

remember, in this sense, the existence of languages with a large number of speakers, but with little presence on the Internet, such as Bengali, Vietnamese, and Iranian Persian. It is important to point out that the work is not limited to endangered languages, the concept of ULCP also includes languages with a large number of speakers, but with few publications on the web, a situation that is also of sociolinguistic interest. Computational products include commercial products, but also include cultural or academic products, such as literature, videos, and corpora for research purposes. The diversification of the web has made it difficult to determine the nature of the offer when it comes to approaching the subject with a linguistic vision, since, as we have already pointed out, the differences between web products tend to blur, and access is often private or restricted.

The questions (15)–(30) pose a profile of the Language Problem with respect to computational products to offer solutions based on language necessities and sensitivities. It is important to delve a little deeper into each of the questions. Thus, although it is true that the answer to (15) is affirmative, it is qualified by an apparent decline of English in the percentage of its audience (share), by dropping significantly to 60%; however, it cannot be affirmed, in any way, that there is a lack of relevant information in that language. Consequently, the question in (16) remains open; a possible investigation to answer it should be posed in terms of the increase of the web and its characteristics (servers, applications, pages, discussion groups, etc.) and in what proportion the products are created in English and in other languages; most probably, we will find that the share corresponding to English decreases by natural effect and has a stable floor. These preliminary questions are closely related to LPP, since it is necessary to answer them in any NLP initiative, among other reasons, to produce successful projects that are sensitive to the needs of the linguistic communities with which it is planned to work.

Questions (17), (18), (19), and (20) are fundamental, because, although the answer is, a priori, negative, there may be an overrepresentation of English, given that it is assumed that the offer must be equal in each language, something that, at least, must be demonstrated, and that in practice is impossible, given that not all subjects have the necessary terminology in the ULCP (quantum mechanics or finance, for example), the opposite may also be true. These questions also shape the position and the attitudes we need to assume when researching language issues and computational applications.

Data on language use in web products assume homogeneity of quality; in (21), we raise the issue of product quality, since if the quality is insufficient or if the supply is irrelevant, then those products will be hardly usable. Resource quality assessments have started to take course, for example, in Swahili and Amharic (Gelas et al. 2011). A well-known case is the Cebuano Language Wikipedia (Sinugboanon), which as of 15 October 2021, had more than five million articles on living things, biological species, cities, and communities, i.e., everything that could be easily transcribed from one Wikipedia language to another (Buchholz 2021; Lokhov 2021). Ninety-nine per cent of these articles were written by the *Lsjbot*[7] bot (Wilson 2020; Manuel 2020). The case of Cebuano is not unique; it has also occurred in other languages. It is worth asking whether detailed articles in Cebuano about medicine, art, or engineering are not more important than long lists of cities.

It is important to note that this method applied to the Cebuano Language was evidenced in the press and in blog posts, not in academic studies. To ignore, these sources would be to silence this phenomenon and probably the very situation of the Cebuano with respect to its presence in technological media. This situation is repeated with various ULCP, for example, in American indigenous languages cited below, such as Kaqchikel. Furthermore, the relevance of web page data to NLP lies in the fact that there is a connection between web publications, circulating information, and web applications, language-based or not.

However, we should also not forget that the resources available in each language community are not comparable; now to claim that the Cebuano Wikipedia contains five million articles is impressive, but the articles, while valid, are not necessarily the most relevant from an encyclopaedic or a relevance point of view to a language community to which they are addressed. The Cebuano experience contrasts with the initiative for Kaqchikel and other Mayan languages of Guatemala on the same Wikipedia (Ortiz 2021), which aims to empower members of these linguistic communities to generate content in the country's vernacular languages. Undoubtedly, as a result, Wikipedia pages in these languages will be more relevant than what a *bot* can produce from translations of foreign city or plant names. This initiative has been supported by several instances and even took place, virtually, a public activity, the Qach'ab'äl

---

[7] A bot is a computer agent programmed for a specific task that performs it automatically; *Lsjbot* was intended for the automatic creation of Wikipedia articles.

pa k'amaya'l Encounter—our languages on the Internet,[8] as part of the dissemination and incentive activities. Consequently, we see that the web products offered in various languages differ in terms of quality and relevance.

The same example of Cebuano demonstrates the relevance of the question posed in (22), which points to the availability of useful information and tools for language communities in their own language. The lack of relevant products may lead to the use of certain tools in a different way than they were intended (23); for example, a machine translation system, which is designed, in principle, to translate sentences, may be used as a multilingual dictionary, making its answers less accurate and more easily misleading, something that users are not always aware of. The relevance of the information is a crucial aspect that is taken up in question (24). It refers, in particular, to the quality of the information. Each linguistic community has the need and the right to relevant and quality information in its own language when accessing government instances or private web applications.

Returning to the general data, it is easy to think that the data on the dominant languages hide many users who resort to them for lack of options in their own languages, an element pointed out by questions (25) and (26), a practice also present in languages that occupy the first positions, due to the fact that the offer of information is also unequal between them. We know that not all users are able to read in other languages, let alone produce content, the assimilation of which is a fundamental issue that goes beyond the scope of this chapter. A concrete example related to these questions can be found in the case of Mayan languages in Guatemala, which have a writing system that is only used by a minority of Mayan speakers, although there is a lack of data to give an accurate picture of the situation. The majority of professionals are educated at higher levels in Spanish, although it is true that, depending on the profession (such as medicine), they may have to be trained to express themselves in an indigenous language, which does not convey knowledge but rather exchanges with members of the communities in which they work.

For this reason, in (27) we highlight, once again, the issue of the extent to which language-specific web products meet the requirements of language communities, according to their expectations and needs. On

---

[8] https://cceguatemala.org/archivos/actividades/encuentro-lenguasindigenas.

topics such as health, it is best to find information in one's own language; if a speaker has to resort to a different language, the effectiveness of communication may be compromised. To a large extent, the lack of relevant information is explained by an inequality in the availability of resources with respect to the languages of influence, which create a disadvantageous situation for the ULCP that could produce serious negative effects, for example, in education and health. Finally, questions (28) and (29) close the series by rephrasing the same aspects of (17)–(20), but in the light of the previous questions. Answering them is important in any research and requires a clear idea of the technologies at play in order to better understand their impact on LPP technological processes.

5   FOR LANGUAGE AWARENESS: CONCLUSIONS

A correct linguistic implementation of web products requires both a thorough understanding of the subtleties related to the functioning of a specific language within the communities where it is used, as well as the formal characteristics they possess. It is in this sense that NLP and LPP should consider the contributions of TL as a way to develop linguistically aware and relevant web products. The primary purpose of the Internet is to improve communications and efficiency by providing a range of services.

While it is true that there is an awareness of multilingualism and NLP (Cracking the Language Barrier Federation 2018), there is still a need to consider more deeply the interaction with communities. Not only the fundamentals of a language are not enough, but the use and the expectations of applications within communities are equally important. For this reason, we posed the series of questions (1)–(30), which will help to define the profile of any project and improve the chances of success.

Thus, the development of applications and the simple search for information require the empowerment of the technology by the linguistic communities themselves, for which an appropriate educational process must be defined for each case. The scarcity of resources for many of the ULCP requires concerted development efforts in line with the expectations and needs of the communities. For this reason, we hope that the questions raised in this chapter will allow for a better investment of resources.

At present, it is not advisable to act in a vacuum, there is a vast accumulated scientific literature and experience that allows a better understanding of the processes at work at the technological, social, and linguistic levels. In this sense, the understanding, analysis, and description of languages, with consideration of their social dimensions, will serve as a bridge between technology, communities, engineering, and science for the benefit of linguistic ecology, but, above all, for a more effective communication of relevant information. It is, therefore, a proposal for linguistic equity that may well be reflected in standardisation and empowerment initiatives, based on the legitimate interests of the sectors involved, which will undoubtedly go beyond linguistics and technological applications in the near future.

References

Academia de las Lenguas Mayas de Guatemala. *K'ULB'IL YOL TWITZ PAXIL*. Academia de las Lenguas Mayas de Guatemala—ALMG. Retrieved 9 June 2022, from https://www.almg.org.gt

Bird, S. (2018). Designing mobile applications for endangered languages. In K. L. Kenneth and L. Campbell (eds.), *The Oxford Handbook of Endangered Languages*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190610029.013.40

Bird, S., Hanke, F. R., Adams, O., and Lee, H. (2014). Aikuma: A mobile app for collaborative language documentation. In *Proceedings of the 2014 Workshop on the Use of Computational Methods in the Study of Endangered Languages* (pp. 1–5). Association for Computational Linguistics. https://doi.org/10.3115/v1/W14-2201

Blokland, R., Fedina, M., Gerstenberger, C., Partanen, N., Riebler, M., and Wilbur, J. (2015). Language documentation meets language technology. In *First International Workshop on Computational Linguistics for Uralic Languages. Septentrio conference series* (pp. 8–18). https://doi.org/10.7557/5.3457

Buchholz, K. (2021). Twenty years after it went online, which are the most popular languages on Wikipedia? *The Wire*. https://thewire.in/tech/wikipedia-popular-languages-english-cebuano

Camacho Caballero, L., and Zevallos Salazar, R. (2020). Computational linguistics for revitalization and polyglotism. *Letras* (Lima), 91 (134), pp. 184–198. https://doi.org/10.30920/letras.91.134.9

Casasola Murillo, E., and Leoni de León, J. A. (2016). Transferring semantic polarity from idiomatic phrases to opinion comments. *Káñina*, 40 (3), pp. 65–76. https://doi.org/10.15517/rk.v40i3.28813

Casasola Murillo, E., Leoni de León, J. A., and Marín Raventós, G. (2017). Evaluation of potential features present in short texts in Spanish in order to classify them by polarity. *CLEI Electronic Journal*, 20 (1). https://doi.org/10.19153/cleiej.20.1.5. https://doi.org/10.19153/cleiej.20.1.5

Chakravarthi, B. R., Arcan, M., and McCrae, J. P. (2018, January). Improving wordnets for under-resourced languages using machine translation. In *Proceedings of the 9th Global Wordnet Conference*, pp. 77–86. Global Wordnet Association, Nanyang Technological University (NTU), Singapore.

Charniak, E. (1984). *Introduction to Artificial Intelligence*. Addison-Wesley.

Constenla Umaña, A. (1992). *Abecedario Ilustrado Bribri*. San José: Comisión Costarricense de Cooperación con la UNESCO.

Cracking the Language Barrier Federation (2018). http://www.cracking-the-language-barrier.eu/

Crystal, David (2006). Chapter 9: English worldwide. In David Denison and Richard M. Hogg (eds.), *A History of the English Language*, pp. 420–439. Cambridge University Press. ISBN 978-0-511-16893-2.

Cunliffe, D., Vlachidis, A., Williams, D., and Tudhope, D. (2021). Natural language processing for under-resourced languages: Developing a Welsh natural language toolkit. *Computer Speech & Language*, p. 101311. https://doi.org/10.1016/j.csl.2021.101311

Eberhard, D. M., Simons, G. F., and Fennig, C. D. (Eds.). (2021). *Ethnologue: How Many Languages Are There in the World?* Twenty-fourth edition. SIL International. https://www.ethnologue.com/guides/how-many-languages

Esselink, B. (2000). *A Practical Guide to Localization*. John Benjamins Pub. Co.

Gambs, S., Killijian, M.-O., and Núñez del Prado Cortez, M. (2014). De-anonymization attack on geolocated data. *Journal of Computer and System Sciences*, 80 (8), pp. 1597–1614. Special Issue on Theory and Applications in Parallel and Distributed Computing Systems. https://doi.org/10.1016/j.jcss.2014.04.024

Gelas, H., Abate, S. T., Besacier, L., and Pellegrino, F. (2011). Quality assessment of crowdsourcing transcriptions for African languages. In *INTERSPEECH, 12th Annual Conference of the International Speech*

Communication Association, pp. 3065–3068. Florence. https://doi.org/10.21437/interspeech.2011-767; https://www.researchgate.net/publication/221478079_Quality_Assessment_of_Crowdsourcing_Transcriptions_for_African_Languages

Goldhahn, D., Sumalvico, M., and Quasthoff, U. (2016). Corpus collection for under-resourced languages with more than one million speakers. In C. Soria et al. (eds.), *CCURL 2016 Collaboration and Computing for Under-Resourced Languages: Towards an Alliance for Digital Language Diversity*, pp. 67–73. LREC. http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-CCURL2016_Proceedings.pdf

Hale, K., Krauss, M., Watahomigie, L. J., Yamamoto, A. Y., Craig, C., Jeanne, L. M., and England, N. C. (1992). Endangered languages. *Language*, 68 (1), pp. 1–42. https://doi.org/10.2307/416368

IBM Cloud Education (2020). *Natural Language Processing (NLP)*

Jackson-Preece, J. (2018). 'Taking Oslo Online': Minority language policy & the internet. In *Language Policy and Conflict Prevention*. Leiden, The Netherlands: Brill | Nijhoff. https://doi.org/10.1163/9789004357754_015

Jurafsky, Daniel, and Martin, James. (2008). *An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, Second Edition. Prentice Hall. https://doi.org/10.1162/coli.b09-001

LaBerge, L., O'Toole, C., Schneider, J., and Smaje, K. (2020, October 5). *How COVID-19 Has Pushed Companies over the Technology Tipping Point-and Transformed Business Forever*. McKinsey & Company. https://doi.org/10.1126/science.abl5437; https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-covid-19-has-pushed-companies-over-the-technology-tipping-point-and-transformed-business-forever

Le, Q. V., and Schuster, M. (2016, September 27). A neural network for machine translation, at production scale. *Google AI Blog*. https://ai.googleblog.com/2016/09/a-neural-network-for-machine.html

Li, B., Erdin, E., Güneş, M. H., Bebis, G., and Shipley, T. (2011). An analysis of anonymizer technology usage. In J. Domingo-Pascual, Y. Shavitt, and S. Uhlig (eds.), *Traffic Monitoring and Analysis. Lecture Notes in Computer Science*, pp. 108–121. Springer. https://doi.org/10.1007/978-3-642-20305-3_10

Llisterri, J. (2003). Lingüística y tecnologías del lenguaje. Lynx. *Panorámica de Estudios Lingüísticos (Departament de Teoria dels Llenguatges, Universitat de València)*, 2, pp. 9–71. Recuperado el 15 de junio de 2022, de http://www.iling.unam.mx/cursocorpus/TecnoLing_Lynx02.pdf

Lokhov, I. (2021). Why are there so many Wikipedia articles in Swedish and Cebuano? *Datawrapper*. https://blog.datawrapper.de/wikipedia-articles-written-by-a-bot/

Manning, C. D., and Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. The MIT Press.

Manuel, J. (2020). Cebuano Wikipedia is 99% written by bots. *Noypigeeks*. https://www.noypigeeks.com/spotlight/cebuano-wikipedia-bots/

Maryland Language Science Center (2018). *Langscape*. University of Maryland. http://langscape.umd.edu/

Nanavati, M., Taylor, N., Aiello, W., and Warfield, A. (2011). Herbert west-deanonymizer. In *6th USENIX Workshop on Hot Topics in Security (HotSec 11)*. San Francisco, CA: USENIX Association.

Ortiz, A. (2021, September 16). The efforts of several digital activists to reach a Wikipedia in indigenous languages. *La Prensa Libre*. https://www.prensa libre.com/vida/tecnologia/los-esfuerzos-de-varios-activistas-digitales-para-lle gar-a-una-wikipedia-en-idiomas-indigenas/

Periñán Pascual, J. C. (2012). En defensa del procesamiento del lenguaje natural fundamentado en la lingüística teórica. *Onomázein: Revista de Lingüística, Filología y Traducción* (26), pp. 13–48. https://doi.org/10.7764/onomaz ein.26.01

Powers, D. M. W., and Turk, C. R. (1989). *Machine Learning of Natural Language*. Springer-Verlag.

Pustejovsky, J., and Stubbs, A. (2013). *Natural Language Annotation for Machine Learning* (First ed.). O'Reilly.

Rumbold, J., and Pierscionek, B. (2018). Contextual anonymization for secondary use of big data in biomedical research: Proposal for an anonymization matrix. *JMIR Medical Informatics*, 6 (4), p. e47. https://doi.org/10.2196/medinform.7096. https://doi.org/10.2196/medinform.7096

Sánchez Avendaño, C. (2013). Lenguas en peligro en Costa Rica: vitalidad, documentación y descripción. *Revista Káñina*, 37 (1), pp. 219–250.https://doi.org/10.15517/rk.v37i1.10589

Stüker, S., Adda, G., Adda-Decker, M., Ambouroue, O., Besacier, L., Blachon, D., Maynard, H., Gauthier, E., Godard, P., Hamlaoui, F., Idiatov, D., Kouarata, G.-N., Lamel, L., Makasso, E.-M., Müller, M., Rialland, A., Van de Velde, M., Yvon, F., and Zerbian, S. (2016). Innovative technologies for under-resourced language documentation: The BULB Project. In C. Soria et al. (Eds.), *CCURL 2016 Collaboration and Computing for Under-Resourced Languages: Towards an Alliance for Digital Language Diversity*, pp. 59–66. Slovenia, LREC. https://doi.org/10.1016/j.procs.2016.04.023; http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-CCURL2016_Proceedings.pdf

W3Techs (2021). *Usage statistics of content languages for websites.* https://w3t echs.com/technologies/overview/content_language

Wilson, Jack (1982). El alfabeto bribri. En E. Margery Peña, *Diccionario fraseológico bribri-español español-bribri*. San José: Editorial de la Universidad de Costa Rica.

Wilson, K. (2020). The world's second largest Wikipedia is written Almost Entirely by One Bot. *Vice*. https://www.vice.com/en/article/4agamm/the-worlds-second-largest-wikipedia-is-written-almost-entirely-by-one-bot

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M., and Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *CoRR*, abs/1609.08144. https://research.google/pubs/pub45610/

Yang, B. (n.d.). *6 common features of top 250 YouTube channels*. Twinword. https://www.twinword.com/blog/features-of-top-250-youtube-channels/

Young, H. (n.d.). The digital language divide. *The Guardian*. http://labs.the guardian.com/digital-language-divide/

Conclusions

Abstract This chapter concludes the book by pointing out five issues that deserve more attention. The first issue is the necessity of researchers leveraging the epistemological and theoretical diversity within the field. The second issue is the study of the conditions under which LPP is effectively able to influence language change, and the context in which such change occurs. The authors argue that more attention should be paid to the effects of the increasing role of the public sector in contemporary societies. The third issue calls for researchers in the field of LPP to be clear in their epistemological and methodological standpoints. The fourth issue is the recent massive use of information technologies for language-based applications, whose current proportions we are just beginning to estimate. A final issue is the difference between scholarship in and about LPP and actual, real-world LPP writing and implementation. A general question remains open: the identification of the epistemological limits of the field of LPP.

Keywords Language policy · Language planning · Epistemology · Theory · Language change · Social policy

This book offers an introduction to some important theoretical and epistemological approaches in the research field of language policy and planning, while presenting some examples of application. Its aim, therefore, is not to provide an exhaustive overview of the diverse and fragmented field of research in LPP, nor to present a new theory. Rather, the aim is to stimulate our readers, in particular undergraduate and postgraduate students, to engage with the theoretical and epistemological diversity in LPP. In this spirit, we would like to conclude this book by pointing out five issues that deserve more attention in future research.

The first issue is the necessity of researchers leveraging the epistemological and theoretical diversity within the field to study, advocate, and consult on language policy and planning activities. Such interdisciplinarity has been a hallmark of the field and it continues to prove beneficial. While the field began as a specialisation within linguistics and sociolinguistics, it has since expanded as a focus within economics, education, anthropology, policy studies, political theory, and legal scholarship, among others. Halliday (1990) argues that "transdisciplinarity" is preferable to "interdisciplinarity" because it engenders new forms of activity that supersede disciplines to address themes, challenges, or problems in applied linguistics. Taken together, the chapters in this book suggest a path towards a transdisciplinary approach. As the field continues to ponder disciplinary diversity, we ask: how do we shift from making connections across disciplines to superseding disciplinary boundaries? A transdisciplinary approach can illuminate (1) the hegemonic impacts of monolingual language policies in schools and society, (2) the potential emancipatory impacts of multilingual language policies in schools and society, and (3) the tension between the ideological structures embedded within language policies and the agency of language policy arbiters, who interpret and appropriate policies in potentially creative and unpredictable ways.

A second issue that deserves more attention is the study of the conditions under which LPP is effectively able to influence language change, and the context in which such change occurs. At its core, LPP starts from the premise that deliberate collective action can modify language change, i.e., the evolution of the status, the acquisition, and corpus of a language. This view is not shared by all in the language sciences (recall that LPP is a rather peripheral area of research in the broad field of linguistics, and it is often considered to be a branch of macro-sociolinguistics, see Hornberger and McKay 2010). As Crystal notes, "many linguists have held

the view that language change is a natural, spontaneous phenomenon, the result of underlying social and/or linguistic forces that it is impossible or undesirable to tamper with. We should 'leave our language alone'" (Crystal 2010: 366). The question that different researchers in LPP have raised, however, is whether it is really possible to "leave our language alone" (Kymlicka 1995; De Schutter 2007; Robichaud 2011; Gazzola 2014. See also Gazzola, this volume). First, all countries make choices about which languages the state apparatus uses for its functioning, spanning from the public administration to courts, from healthcare to public education. Provisions concerning language status, for instance, are contained in the constitutions of 125 of approximately 200 sovereign states (Marten 2016: 76). Language choices, however, can be implicit, and result in de facto language policies, e.g., using one or more dominant languages in the various spheres of government intervention even in the absence of explicit norms. Implicit language policies too can influence language change, in particular the vitality of minority languages, e.g., by expanding or restricting their domains of use in all spheres of governmental action. Language change never occurs in a vacuum, but always in a social and political environment that is characterised and influenced (also) by explicit or implicit language policies resulting from direct or indirect governmental action.

At the same time, more attention in LPP studies should be paid to the effects of the increasing role of the public sector in our lives. The gradual expansion of the roles and tasks of government in the various spheres of collective living is one of the central characteristics of contemporary societies. The government (at any level, national, regional, local) is not only responsible for the minimal tasks that any collective organisation must implement, i.e., the registry and tax office, the administration of justice, the army, and the enforcement of law through police force. The government has many other tasks to fulfil in the modern 'welfare state' including the provision of public education, whose role in language policy and planning is crucial, healthcare, social assistance, and public employment services, to mention the most important ones (in some of these areas, agency can comprise not only the government but also private actors, and the church). For example, in 2019, before the Covid-19 pandemic, the general government expenditures in OECD countries (Organisation for Economic Cooperation and Development) averaged 40.8% of the national Gross Domestic Product (GDP), ranging from 24.2% in Ireland to 55.4% in France (OECD 2022). The most important

item is social protection; that is, pensions, sickness and disability bene-
fits, and unemployment benefits (13.3% of GDP), followed by healthcare
(7.9%). Further, OECD countries spend on average 5.1% of their GDP on
education. It is hard to defend the point of view that government choices,
no matter whether explicit or implicit, as to what languages to use in all
these domains of social policy and in the delivery of collective goods do
not have influence on language change and linguistic vitality. Language
policy for the support of minority languages, therefore, cannot be sepa-
rated from the study of language use in the various areas of government
intervention, and thus from the study of how to protect and promote
minority languages *in* and *through* the structure of the public sector as
a whole. As language policy can sometimes be a transversal dimension of
social policy, more attention should be paid in LPP research to the study
of how social policies work in the first place.

A third issue calls for researchers in the field of LPP to be clear in
their epistemological and methodological standpoints to ensure trustwor-
thiness, and to avoid or at least mitigate biases. While not generalisable,
case study analyses can then be used as a source of comparison with
other situations, if theoretical frameworks behind the empirical data are
explicitly stated. Moreover, even if a case study can focus on a partic-
ular language, all the languages involved in the given context should be
considered, especially in the phase of data collection. Multilingualism is
the default situation in human societies; therefore, speakers have linguistic
repertoires, which include a diversity of language varieties. As a result, if
they choose to use one language in a domain, they are also choosing not
to use all the other languages they have at their disposal. For this reason,
case studies should consider the position of the language targeted in the
LPP action in the larger context of all languages involved.

A fourth issue is the massive use of information technologies for
language-based applications is a recent phenomenon, whose current
proportions we are just beginning to estimate. Only a fraction of the
total number of languages in the world experiences such growth, while
most of the rest of them remain virtually invisible to the information
network. Therefore, we need to appeal to new methodologies and tools
that take into account the contributions of theoretical and field linguistics
in order to approach this reality. A good comprehension of the nature
and challenges of the emerging technologies is becoming a fundamental
knowledge. In addition, the influence of the offer of language-based
products and the definition of standards and protocols by private actors

(in particular the GAFAM[1]) are some elements that we must take into account and that are pending tasks. The creation of resources has a significant impact on the prestige of languages. The concentration of linguistic products and resources in a small number of languages has an impact that we still need to study. Finally, it is important to remember that leaving the initiative in private hands implies a loss of control of an essential element of all cultural heritage by the linguistic communities.

A final issue that deserves attention is the potential differences between scholarship in and about LPP and actual, real-world LPP writing and implementation. It is desirable for LPP practice to be informed by sound theoretical and evidence-based research. At the same time, however, it is important to recognise that LPP practitioners (or policymakers) can be informed by their own respectable epistemological theories, approaches to change and innovation in language, and constraints. The acknowledgement by scholars that practitioner perspectives are not necessarily a simple technical application of pre-defined instructions, but something guided and informed by theory, can enrich our understanding of LPP as a policy activity.

A general question that remains open is the identification of the epistemological limits of the field of LPP. As we have seen in the introduction, contemporary LPP research is both multidisciplinary and interdisciplinary. This leads to a multiplicity of theoretical bases and approaches, which is a rich source of epistemological diversity, but at the same time leads to the paradox that there is no clear answer to the question about the methodological and theoretical foundations of LPP. This is not necessarily a shortcoming, as long as researchers are clear about their epistemological foundation and the field encourages dialogue, both for deontological reasons and for clarity and usability of its results. We hope that this volume has illustrated that different methods, techniques, and approaches within potentially distinct disciplines can start from distant epistemological positions, but they can also achieve a transdisciplinary meeting to develop the field.

---

[1] GAFAM is an acronym for "Google, Apple, Facebook, Amazon, Microsoft", it implies all the web giants.

References

Crystal, David (2010). *The Cambridge Encyclopedia of Language*. 3rd edition. Cambridge: Cambridge University Press.

De Schutter, Helder (2007). Language policy and political philosophy: On the emerging linguistic justice debate. *Language Problems & Language Planning*, 31 (1), pp. 1–23.

Gazzola, Michele (2014). *The Evaluation of Language Regimes. Theory and Application to Multilingual Patent Organisations*. Amsterdam: John Benjamins.

Gobbo, Federico (2015). Transnational communication in a 'glocalized' world: In search of balanced multilingualism. *Language. Communication Information*, 10, pp. 196–208.

Halliday, M. A. K. (1990). New ways of meaning: The challenge to applied linguistics. *Journal of Applied Linguistics*, 6, pp. 7–36.

Hornberger, Nancy H., and McKay, Sandra Lee (2010). *Sociolinguistics and Language Education*. Bristol: Multilingual Matters. https://doi.org/10.21832/9781847692849

Kymlicka, Will (1995). *Multicultural Citizenship: A New Theory of Minority Rights*. Oxford: Oxford University Press.

Marten, Heiko F. (2016). *Sprach(en)politik*. Tübingen: Narr Franke Attempto Verlag.

OECD (2022). General government spending (indicator). https://doi.org/10.1787/a31cbf4d-en. Accessed on 22 July 2022.

Robichaud, David (2011). Justice et politiques linguistiques: pourquoi les laisser-fairistes devraient exiger des interventions de l'État. *Philosophiques*, 38 (2), pp. 419–438.

Index