

Aberystwyth University

Adroddiad ar ddulliau asesu hyfywedd iaith | Methods of Assessing Linguistic Vitality Report

Royles, Elin; Jones, Rhys; Lewis, Huw

DOI:

[10.20391/605eb48c-05e5-44bc-96c5-83394bce046c](https://doi.org/10.20391/605eb48c-05e5-44bc-96c5-83394bce046c)

Publication date:

2024

Citation for published version (APA):

Royles, E., Jones, R., & Lewis, H. (2024). *Adroddiad ar ddulliau asesu hyfywedd iaith | Methods of Assessing Linguistic Vitality Report*. Prifysgol Aberystwyth | Aberystwyth University. <https://doi.org/10.20391/605eb48c-05e5-44bc-96c5-83394bce046c>

Document License

CC BY

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400

email: is@aber.ac.uk



Canolfan Gwleidyddiaeth a Chymdeithas Cymru
Centre for Welsh Politics and Society

Methods of assessing linguistic vitality report

Elin Royles, Rhys Jones and Huw Lewis

Table of Contents

| | |
|---|-----------|
| Recommendations | 3 |
| A: Recommendations for regional or minority languages in general | 3 |
| B: Recommendations specifically for the Welsh language | 4 |
| 1.0 Introduction | 6 |
| 2. Investigating linguistic vitality | 8 |
| 2.1 Defining linguistic vitality..... | 8 |
| 2.2 Assessing linguistic vitality: key factors | 9 |
| Table 1: Categories of factors that influence linguistic vitality | 9 |
| 3.0 Developing indicators for assessing language vitality | 14 |
| 3.1 Good practice and useful international examples | 14 |
| 3.2. Catalonia | 14 |
| Table 2: Catalan Linguistic Indicators System | 16 |
| 3.3 The Basque Country..... | 17 |
| Table 3: The Basque Linguistic Indicator System | 18 |
| 4.0 Data | 20 |
| 4.1 Existing data sources in Wales | 20 |
| 4.2 Linguistic data collection needs in Wales | 22 |
| 4.3 Good practice and useful examples internationally | 24 |
| 4.4 The Basque Country..... | 25 |
| 4.5 General points | 26 |
| 5.0 Understanding and analysing language indicators | 27 |
| 5.1 Technology and data mapping..... | 27 |
| 5.2 Good practice and useful examples internationally | 29 |
| Bibliography | 31 |
| Appendix 1: Information on those consulted during the course of this project | 34 |

Please reference this report as follows:

Elin Royles, Rhys Jones, Huw Lewis (2024) *Methods of assessing linguistic vitality report*. Available in Welsh.

Recommendations

Based on research into ways of assessing the vitality of regional or minority languages, into the value of language indicator systems, and into the potential of mapping methods such as Geographical Information Systems, this report on methods of assessing linguistic vitality assessment presents two sets of recommendations.

A: Recommendations for regional or minority languages in general

Recommendation 1. The report draws attention to the diversity in the range and quality of data gathered to understand the situation of regional or minority languages. Examples of good practice regarding data collection are discussed (Section 4). It would be beneficial for authorities responsible for such languages to assess the extent to which data is collected and the quality of the data collected to understand the vitality of their language/s.

Recommendation 2. The report also draws attention to the challenge of gathering good quality data that covers, not only levels of linguistic ability, but also characteristics such as language use or language transfer (Section 4.3-4.5). We encourage different authorities to consider the report's findings in terms of data collection practices and methods, for instance, in terms of language use across quantitative and qualitative data.

Recommendation 3. The report highlights the value of a system of language indicators that combines various data sources in a systematic way to understand the situation of a language (Section 3). In order to realise this, the report has outlined a framework summarising the main categories of factors that can influence linguistic vitality to be used as a starting point when developing a practical method of assessing the vitality of a specific language (Section 2.2.3 and 2.2.10).

- When developing a system of indicators, the framework should be tailored to emphasise the relevant factors that influence the vitality of a language in the specific context of the language in question (Section 2.2.5).
- When considering the framework and constructing a language indicator system, it could also be tailored to enable understanding the situation of a language at a macro level (e.g. national) and to understand differences in the situation of a language at a more micro/local level. For example, it would be beneficial to be able to assess linguistic vitality taking into account the specific circumstances of areas with a higher density of speakers or in an urban context with a different demographic situation (Section 2.2.6).
- The need to recognise and be alert to the current challenges facing regional and minority languages also needs to be stressed in developing a system of language indicators. This can be achieved through a review of the factors that could be affected by economic or social changes with consequences for the relevance of different elements (Section 2.2.7, 2.2.8).

Recommendation 4. Based on the examples of good practice in language indicator systems discussed, we recommend that the language indicator system and the associated data should be publicly available to different stakeholders, including the public (Section 4).

Recommendation 5. Contributors to the research felt that there was value in developing mapping methods such as Geographical Information Systems to map linguistic vitality (Section 5.2). We believe that there is an opportunity to collaborate across different languages, especially due to the capacity development challenges and the need to share expertise in the field.

Recommendation 6. Alongside the potential of Geographical Information Systems, there is also scope for research and collaboration in order to understand the potential of artificial intelligence to assist in analysing the vitality of regional or minority languages. For example, understanding its potential to facilitate the work of combining relevant administrative data for understanding linguistic vitality. Again, such developments would benefit from collaboration and sharing across different languages.

B: Recommendations specifically for the Welsh language

Due to the specific interest of the research in the assessing the vitality of the Welsh language, the following recommendations are made.

Recommendation 7. Together, the discussion of current data sources in Wales (Section 4.1), the analysis of linguistic data collection needs in Wales (Section 4.2) and discussion of international good practice (Section 4.3) illustrate that there are several data gaps in the relation to the Welsh language. Whatever changes take place in relation to the Census, it is imperative that the data base regarding the Welsh language is not weakened further.

Recommendation 8. Based on the findings (Section 4), there is a need for more robust data that is collected regularly in relation to the Welsh language. In order to realise this, there is a need to:

- Assess whether there are ways to improve the quality of the data already collected.
- Assess whether there are ways to create greater consistency and standardisation across different types of administrative data collection that would be beneficial for assessing linguistic vitality.
- Consider whether it is possible to strengthen the data collection arrangements to be more consistent in order to develop more time series analyses in relation to the Welsh language.
- Respond to the urgent need for quality data regarding language use that is collected in a systematic way and regularly.
- Investigate ways of strengthening the data bases to understand the situation in Wales at a micro/local level.

- Ensure that data is shared and publicly available so that different partners can make use of the data.
- Evaluate the potential of citizen science as a means of filling data gaps.

Recommendation 9. Considering how other leading examples of language revitalisation use language indicator systems to assess linguistic vitality in a way that informs language planning (Section 3), we recommend developing a language indicator system for the Welsh language in order to collate data and to develop a comprehensive picture of linguistic vitality.

A range of situations can be envisaged at different geographical scales, from the local, to the regional, to the national which would benefit from being able to consider the evidence from data associated with a system of language indicators in order for language planning to be conducted in a more purposeful manner. Specific examples are drawing up local development plans or introducing additional interventions to support a minority language in a region experiencing rapid decline.

As a result, further work would be needed to evaluate and develop a report of what are considered to be key indicators for linguistic vitality at different geographical scales, national, regional and community, and in different types of local contexts in Wales.

Recommendation 10. Following developments in the use of Geographical Information Systems for mapping different data in Wales (Section 5.1), a more ambitious step would be to investigate further the potential to combine language data with other relevant data. The intention would be to try to develop a more sophisticated method of analysing linguistic vitality and to pay attention to the relationship between data regarding a language and other variables in order to better understand the implications for linguistic vitality. This would require creating a more quantitative understanding of the relationship between different language variables based on an understanding of the factors that affect linguistic vitality (Section 2.2) and the effect of other variables on linguistic vitality. The aim would be to use this information in order to predict future levels of linguistic vitality, or to assist with decisions regarding targeting investment.

1.0 Introduction

1.1 This report outlines the main findings of a research project examining ways of developing alternative means of understanding the vitality of regional or minority languages such as Welsh. As understanding the degree of linguistic vitality and fragility is central to language planning and efforts to revive regional or minority languages, a number of initiatives have sought to assess linguistic vitality and the need for further development of the field is also recognised. Notably, Professor Colin H. Williams (2022) emphasised the potential of using mapping methods such as Geographical Information Systems. Consequently, this work aims to contribute to assessing the value of developing methods of assessing linguistic vitality for regional or minority languages generally, and specifically for the Welsh language in Wales.

1.2 As a result of initial research in the field by the authors, the work associated with this report was commissioned by the Welsh Language Commissioner in order to specifically examine key elements in analysing linguistic vitality, namely defining linguistic vitality, indicators, data, the potential of technology and data mapping, and international good practice. We are very grateful to the Commissioner for supporting this work.

1.3 The various activities that feed into the research were as follows:

- An online seminar with experts on minority languages from several countries, and those who have been involved in the development of linguistic indicator systems;
- An in-person seminar with practitioners from Wales involved in collecting data associated with understanding linguistic vitality and who make use of this data in their work;
- Meeting with officials from the Secretaria de Política Lingüística, Generalitat de Catalunya responsible for language indicators and reports on the Catalan language;
- Meetings with relevant Welsh Government officials.

A list of contributors to the research is available in Appendix 1. We are very grateful to everyone who contributed to the project.

1.4 The report provides a summary of the main findings, and we also outline areas where further work is required. First, we discuss how the concept of language vitality is defined in the relevant research literature and consider the type of factors that tend to be listed as relevant to assessing the vitality of a regional or minority language such as Welsh (Part 2). Then, the process of identifying specific indicators that can be used to assess linguistic vitality is discussed (Part 3). The type of data sources currently available in Wales that could be incorporated into an assessment of Welsh language vitality are then reviewed, taking into account examples and international good practice of collating this type of data (Part 4). In parts 3 and 4, we

focus on the Catalan and Basque Country cases as examples of international good practice in this respect. They are also prominent in their efforts to revive a minority or regional language as in the case with the Welsh language in Wales. The final part of the report (Part 5) focuses on ways of understanding and analysing indicators and data regarding linguistic vitality and assesses the potential of mapping methods and technology to analyse the vitality of regional or minority languages. The Recommendations at the beginning of the report (pp.3–5) are organised according to i) points that are generally relevant to regional or minority languages and ii) points that are relevant to the case of the Welsh language in Wales.

- 1.5 This project has sought to be relevant to various regional or minority languages in Europe, and particularly to the context of the Welsh language and other regional languages in the United Kingdom as the Office for National Statistics considers changes to the census arrangements, the main source of linguistic data in this case.

2. Investigating linguistic vitality

2.1 Defining linguistic vitality

2.1.1 Vitality is a concept that has claimed a central place in discussions seeking to analyse the condition of different languages. In general, there is a tendency to assume that a language with a high level of vitality is more likely to survive over time and to continue to be used. For example, in an early example of a study using the concept, Giles *et al* argued: 'The more vitality a linguistic group has, the more likely it will survive and thrive as a collective entity in an intergroup context' (Giles, Bourhis and Taylor 1977: 308).

2.1.2 As a result, in efforts to maintain or revive languages, it has been argued that analysing the level of vitality of the language in question is a key step that will influence the type of language planning interventions that should be prioritised.

2.1.3 Linguistic vitality tends to be used as a term encompassing a range of possible circumstances, rather than an adjective that describes one specific linguistic condition (Tsunoda 2005). This is evident as the research literature associated with the subject includes a wide variety of scales seeking to describe the various levels of vitality that can characterise different language communities. For example, Grenoble and Whaley (2006: 18) propose the following five-level scale:

| | | | | |
|------|-----------|--------------|----------|------|
| Safe | In danger | Disappearing | Moribund | Dead |
|------|-----------|--------------|----------|------|

2.1.4 Following from this, it is necessary to distinguish between 'linguistic vitality' and the related notion of 'viable language'. The latter is an adjective used to describe a specific linguistic condition and is often contrasted with other specific adjectives, for example a 'fragile language' or a 'dead language' (Tsunoda 2005: 9–13). However, scrutinising the research literature illustrates the different interpretations of what can be described as a condition of viability. Some will use viable as an adjective to describe a relatively ideal situation where the language has a significant number of speakers and those use the language regularly in all aspects of their life (e.g. Grenoble and Whaley 2006). From this perspective, viable is considered as a description synonymous with flourishing. Others will use it to describe more limited circumstances where the language is not necessarily used in all spheres of life but has a stable core of speakers who use it regularly in certain contexts and that they continue to transmit it to the next generation (e.g. Fishman 1991). From this perspective, although the language is not ubiquitous, and therefore not necessarily thriving, it can be described as viable as it is likely to survive into the future.

2.2 Assessing linguistic vitality: key factors

2.2.1 There is recognition that developing a balanced analysis of the level of vitality of different languages is not easy given the need to consider a range of interconnected factors. For example, the authors of the *Language Vitality and Endangerment* report published by UNESCO (2003: 7) emphasised: 'No single factor alone can be used to assess a language's vitality ... language communities are complex and diverse'.

2.2.2 To address this challenge, a range of researchers specialising in subjects such as language shift and language revitalisation have developed frameworks seeking to summarise the range of key factors that should be taken into account when evaluating the vitality of different languages (e.g. Edwards 1992; Fishman 1991; Giles, Bourhis and Taylor 1977; Lewis and Simons 2010, 2016; Tsunoda 2005; UNESCO 2003; see also Baker and Wright 2021 for a useful summary of this literature).

2.2.3 For the purpose of this report, a review of the research literature was undertaken, giving specific consideration to the context of European regional or minority languages. On the basis of this review and taking into account the findings of previous research (Lewis and Royles 2018; Jones and Lewis 2019; Lewis and McLeod 2021), a composite framework has been developed summarising the main categories of factors that tend to be considered as influencing linguistic vitality, along with the type of specific characteristics that tend to be associated with each category. The main categories are outlined in Table 1 below.

Table 1: Categories of factors that influence linguistic vitality

| | |
|-----------------------------------|--|
| Demographics | The number of speakers and their distribution across the territory and across the generations. |
| Acquisition practices | How and when the language was learnt by its speakers and the level of competence. |
| Language use practices | The use of the language across different social domains, whether in private/informal domains or official/formal ones, and whether the domains are physical or virtual. |
| Institutional support | The recognition and support extended to the language by government, along with other public and private institutions. |
| Psychology | Attitudes towards the language held by different groups, together with the status or prestige attributed to it. |
| Linguistic characteristics | The condition of the language itself, e.g. to what extent it has been standardised, and whether new terminology is being developed etc. |
| Economic context | The economic status of language speakers and the nature of the socio-economic circumstances that they face. |

2.2.4 The above framework and the general categories set out could be used as a starting point when considering how to develop a practical method for assessing the vitality of a specific language. However, we stress that the framework is a starting point. Based on our research, together with the workshop discussions held as part of the project, attention was drawn to various further points that call for consideration when seeking to apply the framework to specific cases. Some of these points are discussed below (2.2.5–2.2.8).

2.2.5 Although the research literature tends to develop frameworks of key categories that can be applied in order to assess vitality across a wide range of cases, contributors to this project argued that it is important to give further consideration to the specific context of the language in question. Despite limiting the focus to European regional or minority languages, the range of linguistic situations that can arise varies substantially (e.g. in terms of demographics, institutional support, usage habits etc.). As a result, the significance of a number of the general categories listed above to the process of assessing vitality can vary according to the exact linguistic context under study. This has important implications if there is an intention to move from identifying general themes (as in this section) to developing indicators that can be measured based on data (see Section 3 below).

2.2.6 Another consideration is the extent to which greater consideration needs to be given to the geographical context when trying to decide which factors should be taken into account when trying to assess the vitality of a specific language. It was noted that the level of vitality conveyed by considering a collection of factors can vary according to the geographical level in question (e.g. the picture can appear relatively healthy at a local level but less positive when viewed at a national level). In addition, it was suggested that different factors may need to be considered when assessing vitality in an urban context compared to a more rural context, due to the different life patterns in both types of location.

2.2.7 It was also stressed that the process of identifying actors for assessing the vitality of a language should be reviewed, for example every decade, as changes in the nature of the economy or society can affect the relevance of different elements or concepts (Lewis and Royles 2018; Lewis and McLeod 2021). For example, it was noted that the research literature has tended to view the process by which children acquire language as one that occurs primarily in the family home, guided by two parents. However, the greater diversity in the nature of households today needs to be recognised, and the fact that changes in work and childcare practices mean that language acquisition processes during childhood can be more multifaceted. In addition, when considering language use it was noted that there is a tendency in the traditional research literature to emphasise religion as a key domain, whereas 'new' online domains tend to be less prominent, although language practices in these contexts are likely to be more significant today.

2.2.8 Building on the last point, there was a discussion about how best to recognise the increasing but multifaceted influence of digital technology and new media when

constructing a framework of factors that influence linguistic vitality. On the one hand, it was noted that such developments contribute to the creation of new language domains and that it is necessary to assess the extent to which a specific language is used in these domains when assessing its vitality (Cunliffe 2021). On the other hand, it was noted that the growing influence of technology (including more recent developments such as artificial intelligence) can also affect the situation of regional or minority languages in more fundamental ways. This is due to its transformative impact on the way people live and interact with each other - in this context it is a factor more akin to urbanisation or mobility in earlier periods.

2.2.9 While recognising the need to consider the above points when seeking to apply the framework outlined in Table 1 in specific cases, it is worth noting that the general relevance of the different categories was not questioned by those who contributed to the project workshops. This therefore strengthens our conclusion that the framework can be used as a suitable starting point when trying to develop a practical method for assessing the vitality of a specific language.

2.2.10 Taking that into account, below we expand upon the content of Table 1 by outlining the type of specific characteristics that tend to be associated with each category:

- **Demographics:** numbers of speakers and their distribution across the territory and across generations:

- Absolute numbers of speakers
- Speakers of the language as a percentage of the population
- The density of speakers in a specific area / specific areas
- The age profile of the speakers (i.e. distribution across generations, or whether intergenerational transfer is taking place)
- Endogamy/exogamy trends (or the linguistic nature of households)
- Migration trends (in-migration/out-migration)

- **Acquisition practices:** how and when the language was learned by its speakers (e.g. how the language is passed from one generation to another) and the level of achievement:

- Learning via family members
- Learning through pre-school education
- Learning through statutory education
- Learning as an adult
- The level of literacy and fluency of the speakers

- **Language use practices:** the use of the language across different social domains, whether in private/informal domains or official/formal ones, and whether the domains are physical or virtual:

- The family/home
- The community/neighbourhood

- Education (different levels)
- Workplace
- Public services
- World of commerce (shopping, banking etc.)
- Social life/leisure
- Online and on social media

• **Institutional support:** the extent to which language use and the process of creating speakers are supported and facilitated by the policies and work practices of key organisations (organisations that provide services, information, goods or opportunities leisure and also employers):

- Government bodies (local, regional, national, state)
- Public sector bodies
- The private sector
- Third sector organisations
- Civil society
- The media

• **Psychology:** different people's attitudes towards the language, together with the status or prestige attributed to it:

- The attitudes of those that speak the language
- Attitudes of speakers of other languages within society
- The prestige of the language in official contexts
- The prestige of the language in economic contexts (i.e. the market)
- The prestige of the language in socio-cultural contexts (i.e. is the language attractive or 'cool')

• **Linguistic characteristics:** different aspects of the condition of the language itself:

- Standardisation – is there one standard form of the language that stands above regional or dialect divisions?
- Modernization (modernisation) - what is the length and breadth of the language's vocabulary and terminology?

• **Socio-economic context:** the economic status of persons speaking the language, and the nature of the socio-economic circumstances faced:

- The economic status of the speakers of the language (wealth/professional status/level of economic activity/unemployment)
- Employment opportunities – in the territory in general/in higher density areas in particular
- Availability of suitable/affordable homes - in the territory in general/in higher density areas in particular

- How attractive is life – in the territory in general/in higher density areas in particular.

2.2.11 Before concluding this section we note one important limitation relating to the framework outlined above. While the framework summarises the main categories of factors which tend to be viewed as ones that influence linguistic vitality, it does not offer much guidance as to how the interrelationship between some of these factors should be interpreted, or the extent to which greater emphasis should be placed on certain factors. This echoes criticism previously voiced by Colin Baker (2011: 75): 'A list of the relative importance of these factors is simplistic because the factors interact and intermingle in a complicated equation. Such a list does not distinguish the more important factors in language shift. Nor does it reveal the processes and mechanisms of language shift'.

2.2.12 This point was considered when reviewing the academic literature and it was also discussed with different interviewees who contributed to the research. However, it was not possible to find current methods that could be applied in order to go beyond listing factors and to start considering the relationship between them and their relative weighting. As a result, further research is required in order to tackle this key challenge.

2.2.13 Finally, although the research literature enables the establishment of a general understanding of the type of factors that can theoretically influence the vitality of a language, there is a need for further refinement to develop a practical method of assessing vitality in a specific context. A step in this process is to define a series of specific indicators. To achieve their purpose, these indicators will need to be ones that can be measured based on data that can be collected on a regular basis. The next sections will turn to consider this task further.

3.0 Developing indicators for assessing language vitality

3.1 Good practice and useful international examples

3.1.1 Gazzola and Iannàccaro (2024: 331) note that the use of linguistic indicators in language planning and policy in general is a relatively recent development compared to other policy areas and is not very structured to date. In the context of regional or minority languages, indicators are predominantly used for two main purposes:

- to assess the current situation of a minority language;
- to monitor the implementation and effectiveness of interventions that seek to promote a minority language, e.g. implementation of a language strategy.

3.1.2 Although there is sometimes an interconnection between the two purposes, e.g. that specific parts of a system of indicators developed to assess linguistic vitality are used to monitor the implementation of a language strategy, the main focus of this research is systems to assess the vitality of regional or minority languages. In the case of Catalan and Basque we see developed examples of linguistic indicator systems, outlined in turn below.

3.2. Catalonia

El Sistema d'Indicadors Lingüístics (SIL) is the System of Linguistic Indicators used in Catalonia in relation to the Catalan language. This system was first introduced in 2006–08 and reinstated from 2014 onwards.

3.2.1 When constructing a system of linguistic indicators in Catalonia, language indicators were considered as measurements that would provide relevant information about the linguistic situation of a body, its environment, or society in general, helping to shape political decisions and language planning processes.¹

3.2.2 In terms of the process, the points identified by the Commission of Experts as necessary when creating an indicator system in the first place are helpful:

- **Definitional clarity:** Indicators that would be easy to understand avoiding ambiguity or complexity.
- **Relevance:** Emphasis on being able to measure the most relevant elements.
- **Feasibility:** The need to be able to collect the information within the resources that are, or will be, available.
- **Stability:** The need to be able to act as serialised data.

¹ Solé Camardons, Joan (2003) 'El Sistema d'Indicadors Lingüístics (SIL): finalitat i característiques generals,' *Revista de Sociolingüística*.

- **Reliability:** Any change in indicators needs to match what is being measured.
- **Clarity:** The need to be comprehensible to non-experts.²

3.2.3 Emphasis was also placed on indicators that would be in line with the following recommendations:

- The interest of the citizen - citizens should be able to understand the usefulness of the linguistic indicators and the need to link between what is of interest to citizens and to experts.
- Need to differentiate between a system to analyse and report on a sociolinguistic situation and a system that evaluates language policy because they have different objectives and methods.
- A public system of indicators where citizens can access the data linked to the indicators.
- The importance of ensuring consistency between survey protocols on questions of language use and identities in order to ensure that language data is available at low cost.
- Specific points about the nature of the data, for example the need to differentiate and collect data on different elements of language use as with collecting data on linguistic ability, the value of data on citizens' subjective impressions of the linguistic situation.

3.2.4 The system is outlined in Table 2. The current version encompasses 8 areas, 6 for Catalonia itself and 2 for locations beyond the Catalan Autonomous Region. The system focuses on indicators relating to the situation of the language in specific contexts and in different parts of society. Indicators mainly relate to a level of understanding of the language, the use of it, as well as the linguistic offer or the services that are provided in a language and the use of them.

3.2.5 The indicator system results are presented on a website and in publications such as the Language Policy Report. In terms of the practical impact of the indicator system on Catalan Government actions, it was emphasised that the data is key to understanding the language situation in practice, and enables responding to those situations. Reference was made to the following examples:

- Contributing to understanding the linguistic situation among those not born in Catalonia that led to developing specific language learning provision;
- Recognising a shortage of doctors who speak Catalan, leading to a specific focus in language policy on doctors learning Catalan;
- Highlighting the situation of young people and their language use on multi-media platforms, leading to the development of a channel for teenagers;
- Surveys in trading locations showed limitations to services in Catalan, leading to offering dedicated sessions for those working in the sector.

² Ibid.

Beyond this, the results are used by civil society, researchers, journalists and they inform public discussion.

Table 2: Catalan Linguistic Indicators System³

| Field | Subfield |
|--|--|
| 1. Population | Linguistic understanding |
| | Mother tongue, the language people identify with, the language they usually use |
| | Language use |
| 2. Public Administration | Citizens' relationship with the administrations |
| | Generalitat administration |
| | Local administration |
| | Administration of justice and the law |
| | State-level administration |
| 3. Education | Primary Education |
| | Secondary Education |
| | Baccalaureate and professional training |
| | University education |
| | Adult education |
| 4. Health and social services | (Including organisations and civic society, clubs and sports clubs, youth activity) |
| 5. Media, cultural consumption and the digital world | The Media (television, radio and the press) |
| | Cultural use (editorial production, cinema, performing arts, music and exhibitions) |
| | The digital world (language in online media, digital media and video games) |
| 6. The socio-economic world | The offer in Catalan or the level of services/provision |
| | The use made of the services/ provision |
| 7. The territories that speak the Catalan language | Linguistic understanding |
| | Mother tongue, the language people identify with, the language that they usually use |
| | Language use |
| | Language of cultural and media engagement in Catalan-speaking territories |
| 8. External visualisation | Visualising the language |

³ Generalitat de Catalunya (2023) SIL, Conceptualització i estructura: https://llengua.gencat.cat/ca/serveis/dades_i_estudis/sil/conceptualitzacio-estructura/.

3.3 The Basque Country

3.3.1 The Basque Linguistic Indicator System is the **Euskaraen Adieralze Sistema (EAS)** launched in 2013 and implemented from 2015–17 onwards. The administrations across the territories where Basque is spoken are partners in this initiative and emphasis is placed on developing indicators regarding the social situation of the language throughout the Basque areas. The current system includes 33 different indicators. Although the first is of a qualitative nature, the rest is more quantitative. Emphasis is placed on understanding the language and its use with indicators organised according to different domains.

3.3.2 The central principles of this system of indicators are that the information published periodically should be in public ownership, that it functions as a public service and is based on the widest political, social and technical consensus. It should be:

- easy to understand
- mainly quantitative, but that some qualitative elements are useful and meaningful
- covers the territories where Basque is spoken
- takes into account different levels of administration and their linguistic responsibilities.

Table 3: The Basque Linguistic Indicator System

| Linguistic Indicator | Sub-dimension | Further sub-dimension |
|--|---|---|
| 1. Legal Status of the Basque Language | | |
| 2. Linguistic Ability, Use and Attitude | Linguistic Competence in Basque | |
| | Use of Basque | |
| | Attitude towards efforts to promote the Basque language | |
| 3. Family, Education and Teaching Basque | Family | Mother tongue |
| | | Language used at home |
| | Education | Compulsory education, secondary education and vocational training |
| | | University |
| | Teaching Basque for adults | |
| 4. Indicators according to domains | Administration and basic services | |
| | Socio-economic | |
| | Culture | |
| | Media | |
| | Informal relations and leisure provision | |
| | Information technology | |
| 5. Socio-demographic | Data including: Population by age and sex Population by place of origin Births and deaths rate | |

3.3.3 Overall, despite the differences between the language indicator systems in Catalonia and the Basque Country, there are many similarities in terms of the type of indicators included. Based on our research and the discussions, main points in terms of evaluating the different indicator systems are as follows:

3.3.4 It is important to note the emphasis in both systems on indicators relating to language use.

- One difference between the two indicator systems is the lack of reference to demographic indicators in the Catalan system whereas this is a clear feature in the Basque system (Linguistic Indicator 5). This gap is somewhat surprising considering the extent to which migration from other parts of the Spanish state and international migration have affected the Catalan language. We understand that this gap is now recognised in the Catalan system and that ongoing work on indicators attempts to give more attention to demographic aspects.
- As language transfer is such a crucial element to understanding linguistic vitality, questions were raised about the extent to which the two indicator systems are able to understand this as there is no specific indicator on this in both cases. Instead, information about language use within families and among different members of the family is the method of assessing the degree of transfer. It should also be noted that family language transmission in a case like Catalonia is less of a challenge than in other cases.
- Another specific feature is ideology and linguistic attitudes. There was recognition of the challenge of assessing attitudes to language as part of a system of indicators. Both systems seek to understand attitudes by using information regarding issues such as identification with a language and attitudes towards issues such as measures to support the Basque language. At the same time, in the case of Catalonia more emphasis is placed on indicators of language use than on assessing attitudes. The argument was that levels of use are ultimately most important, and they are also an important way of understanding and highlighting the effects of attitudes.
- Considering the side effects of digital developments, both indicator systems pay attention to the use of the language in different online contexts and the availability of online services by accessing information about the provision and use of services from the private sector on digital platforms. In the case of Catalonia, for example, the media and online domain considered as covering the 9 domains of language use that are included in the language indicator system. Contributors from Wales thought that it is important to pay stronger attention to the digital domains as they become increasingly important, e.g. include a specific indicator on this.
- It should be noted that the Catalan system pays attention to the situation of the language beyond the Autonomous Community which raises questions about the extent to which the number of speakers etc beyond the territory should be considered.

4.0 Data

Having established the type of indicators and language indicator systems currently used in the Catalan and Basque cases, we turn to consider the type of data used to understand the current language situation, firstly looking at Wales and then at Catalonia and the Basque Country.

4.1 Existing data sources in Wales

4.1.1 The Census: A number emphasised the strength of the census as a highly established source of data available over a large period of time that enables an understanding of the linguistic situation in Wales at a national level and at other geographical scales, including very local level data at ward level. Another strength is the way it enables linguistic elements to be considered alongside a wide range of other aspects. At the same time, the degree of dependence on this source was stressed alongside the fact that it is conducted every ten years. Consequently, the data becomes dated and the 10-year period between each census means that there are data gaps in the intervening periods.

4.1.2 The Language Use Survey: This is currently the main source to understand language use. The surveys are carried out occasionally (2004–06, 2013–15, 2019–20) and part of the data collection process is linked to the National Survey. The focus is on two age groups 3–15 and 16+. COVID-19 affected data collection for the 2019–20 survey and therefore it was not possible to systematically sample at local authority level as is usually the case (reference is made for example to a systematic sampling of 450 for each local authority in the case of the 2013–15 survey). Consequently, the data is not available at local authority level. As a result, there are questions about the reliability of the detailed data for 2019–20. More general shortcomings of this survey are that it is not carried out regularly or consistently and that it does not assess in greater detail below local authority level.

4.1.3 The National Survey of Wales: This survey is conducted annually with a sample of twelve thousand adults aged 16 and over (organised by the Office for National Statistics 2016–23). Since its introduction in 2012–13, it has asked questions annually regarding linguistic ability. There are also additional questions about the Welsh language in some years. For example, a series of questions were asked about attitudes to the Welsh language in 2017–18, questions about language use in the workplace in 2018–19 and a series of questions about the language individuals would use when attending different locations in 2019–20. Given the size of the sample and the opportunity to ask more detailed additional questions, it has the potential to be a useful source of data. For instance, it was used in the formulation of the *Cymraeg 2050* language strategy. However, beyond the questions about linguistic ability, there is no pattern of asking additional questions regarding the Welsh language in more than one version of the survey. A probable reason for this is the need to compete to include questions on subjects in the survey and a tendency not to repeat the same series of questions from year to year. The effect of

these arrangements is to limit the usefulness of the survey as a source of time series data about the Welsh language and its responsiveness to the need to understand trends over time. Furthermore, local authority level is the lowest level at which the data can be analysed.

4.1.4 Annual Survey of the Population: a survey by the National Statistics Service which also provides data on the number who can speak Welsh. It reports on trends and the local authority level is again the lowest level for analysing the data.

4.1.5 Pupil Level Annual School Census (PLASC): The value of this data on statutory education which is collected annually was also noted as it enables an understanding of the linguistic situation in that age period by recording data on the first language of pupils, different data on the language of education and it records data on issues such as pupils' ethnic background, degrees of poverty and deprivation. Based on this data, the Government has created a data dashboard about the Welsh language in order to support local authorities in developing their Welsh in Education Strategic Plans. It was noted that some elements of the data are of high quality and reliable but that there are generally potential limitations in terms of the reliability of the data due to the nature of the data collection process in relation to certain questions and differences between different data sources, e.g. how often the data is updated.

4.1.6 Reference was made to a range of **other administrative surveys** which collect data relevant to the Welsh language including the Lifelong learning Wales Record (LLWR) (Post-16), a survey conducted by the National Centre for Learning Welsh Survey, Education Workforce Council Data, Care Inspectorate Wales data on childcare and Mudiad Meithrin data on their workforce, and the data in Local Well-being Plans. Some of these data sources are pulled together on Stats Cymru, the Welsh language catalogue.

4.1.7 In the case of Wales, the emphasis is currently on gathering data for monitoring and assessing the implementation of the Welsh language strategy, *Cymraeg 2050*. This provides a means of measuring progress in implementing the strategy and of understanding the contribution of different activities to wider policy objectives. It is assessed alongside the trajectory and targets set out in the work programmes every five years (see Welsh Government, 2022). As a result, many statistics are available in *Cymraeg 2050* annual reports that outline performance and note progress, successes and challenges.

4.1.8 More qualitative research produced regarding the Welsh language provides a greater understanding of the factors that affect language choices and behaviour and give a more detailed picture in different language domains or specific geographical areas. However, there are limitations to this research in terms of the amount of work that is produced due to capacity issues, a small number of researchers on language policy and planning in Wales and issues such as the challenges of securing research funding to support work relating to the Welsh language, as is the case for a number of other regional or minority languages.

4.1.9 In addition to the points stated above about the main sources of data in relation to the Welsh language (4.1.1-4.1.6), general points were outlined in terms of the limitations of the data:

- A lack of evidence which creates a challenge in terms of understanding the impact of migration and mobility on the Welsh language. This is important in order to understand population movements within and outside of Wales and the changes in the number of speakers in different areas. To date, there is no consistent method in place to analyse this.
- The challenge that quantitative data sources often contradict one another, including about basic issues such as people's language ability in Welsh.
- Reflecting some of the points above, general points were made about the quality of the quantitative data collected, the extent to which consistency between the questions asked in different surveys enables comparison, the reliability of the data based on the methods of collection and the evidence base used in order to present responses e.g. risks that questions about language skills are answered on a fairly subjective basis, without sufficient grounds and inconsistently.
- As requirements have already been placed on different bodies to report on data, it was suggested that there was potential to include additional questions that would increase the evidence on linguistic issues. Specific attention would need to be given to the type of questions asked in order to obtain relevant data and a need for consistency across different data.
- Whereas combining different administrative data with linguistic data can be a means of responding to some of the current data challenges, that has its own challenges in terms of resources, people, time and costs. For example, issues such as permission to link data, the systems and processing for sharing and accessing the data, related costs in terms of human and financial resources.
- Implicit in this, linguistic data collection inevitably takes place at different levels within Wales. Functions in terms of decision making and oversight of those processes and allowing access to data also go beyond the borders of Wales due to the functions of the UK-wide Office for National Statistics. To what extent do these governance arrangements affect the ability to collect extensive and purposeful data to assess the vitality of the Welsh language?

4.2 Linguistic data collection needs in Wales

4.2.1 One of the implicit challenges in this discussion in Wales is the current timing. In particular, the proposed changes to the Census are not known and those will have

significant implications. It was also noted that the Education Bill has implications for the categorisation of essential elements.

4.2.2 As a result of the current situation and some of the challenges highlighted, the points made regarding linguistic data collection needs in Wales can be summarised as follows:

- The need to collect data consistently and more regularly was emphasised in order to be able to develop time series analyses of the situation of the Welsh language. Quality data should be collected to facilitate detailed analysis and highlight trends and trajectories in terms of use of the language and its vitality. At the same time, there is a need to be clear as to what the purpose is when collecting data.
- Specific attention was drawn to the weakness of the current arrangements in terms of collecting data about language use and the clear gap considering that the second main target of the *Cymraeg 2050* strategy is increasing the numbers who use the Welsh language. There are no robust and reliable data arrangements in place to assess this. Furthermore, it was noted how crucial it is to develop a better understanding of language use, including at a micro level. Among the suggestions was to use a range of sources to create a more comprehensive picture. Reference was made to the Basque Street surveys as an excellent example in this respect, including to identify trends over time.
- A broader general point is the degree to which different data sources enable linguistic analysis at different geographical scales, including micro levels.
- Among the areas of significance to understanding Welsh linguistic vitality, the following were prioritised: understanding the effects of immigration and migration in Wales, the extent of language transmission within families, the linguistic practices of young people, understanding the relationship between the economy and the language, technology and language use, the private sector and the potential to monitor the use of Welsh in the workplace.
- The potential to use citizen science methods was discussed as a way of gathering more data about the situation of the Welsh language. The weaknesses associated with this method must be weighed up, specifically regarding the reliability of the data, and its strengths, namely the ability to fill data gaps (e.g. in enabling more regular data collection, and spatial gaps in enabling data collection at more sites). Other agencies in Wales (e.g. Natural Resources Wales) make wider use of this method and lessons could be learned from their experience.

4.3 Good practice and useful examples internationally

4.3.1 Here, we discuss in detail the example of Catalonia, before discussing key elements in the Basque Country case where further work could be undertaken. We provide general comments based on the research and discussions.

4.3.2 In Catalonia, the Language Politics Secretariat (Secretaria de Política Lingüística) leads on the language indicators system, although several surveys that feed into it are implemented by the Institute of Catalan Statistics. The Catalan language indicators system tries to bring together in a systematic way a large range of data from different sources. The emphasis is on collecting times series data from official statistics or series studies. The indicator system is updated every year, incorporating the most current data available. Overall, a publication in 2022 indicated that around 40 different sources are used in the system, of which 63.2% are administrative records, 23.7% are surveys, 10.6% are surveys and ethnographic work and 2.6% are summary indicators. In terms of their sources, 86.6% come from public bodies. 79% are updated annually.

4.3.3 The main sources of data used are the following (indicating the last year of data collection):

- Population Language Use Survey (EULP) 2018 > 2023
- Linguistic Census / Statistics on Main Characteristics of the Population and Housing (ECEPH) 2021
- Catalan Education Indicator System (CSASE) 2023
- Socio-demographic and linguistic study of 4th year ESO students in Catalonia (2022)
- Survey of cultural participation in Catalonia (EPCC) 2023
- Cultural Statistics of Catalonia (ICEC) 2023
- Opinion Studies Centre of the Generalitat of Catalonia (CEO) 2024
- Territorial ICT household survey 2021 (2023)
- Survey on the language offer in shops and restaurants (OFERCAT) 2023
- Other data records (Department of Justice, Secretariat of Universities and Research)
- Other studies and data sources: General Media Survey, OJD etc

4.3.4 The cornerstone of the indicator system is the EULP survey held every five years. An online questionnaire is used and the sample of 9 thousand over the age of 15 enables an understanding of the situation according to demographics in detail. The final sample in a recent case was 8,780 completed questionnaires. There is also a geographical element to this sampling by collecting a sample of statistical significance in 16 specific areas, 8 in the metropolitan area where the language has experienced more challenging circumstances, particularly as a result of immigration and where the linguistic situation is similar of being more diverse, 7 other territorial areas, and in Val d'Aran.

4.3.5 Overall, across the data sources, the emphasis is on collecting the most essential and important data in a logical way to understand what is happening in terms of the language in practice, focusing in particular on levels of understanding of the language, and a particular emphasis on language use in different fields and contexts. As part of this, and in line with the indicator system, data is collected on the linguistic offer or the services that are provided in a language and their use in the field of public administration and the socio-economic world. Overall, the range of data used to investigate language use in different situations is impressive. It also includes questions about the respondent's usual language choices in specific circumstances, for example where a respondent speaks Catalan but receives an answer in Castilian.

4.3.6 Due to the range of data sources, triangulation can be achieved by considering responses to the same questions in different surveys. In addition, some selective variables have been identified as key to assessing the credibility of surveys, especially when data findings vary significantly or unexpectedly. In the case of Catalonia, these are place of birth and first language as these elements have such a significant impact on other survey responses.

4.3.7 Consideration had been given to the potential of other data sources, for example information about language use collected by companies e.g. Netflix. It was explained that they had faced difficulties in accessing such data due to the companies' lack of willingness to share.

4.3.8 With regard to the further direction of the Catalan surveys, they aim to expand the areas and domains to understand language use practices. For example, until recently there was no information about language use in restaurants, places of worship/religion, or sports. New questions were added on these.

4.4 The Basque Country

The Basque Country indicator system is considered to be an important method of collecting and systematically organising useful data to understand the situation of the language.

4.4.1 The main data sources of the indicator system which enable understanding the situation of the Basque language in different parts of society at a specific time and the trends over time are the following:

- A Sociolinguistic Survey conducted every five years among people over 15 years of age in the Basque Country (data 1991–2021)
- A Sociolinguistic Map conducted every five years among individuals aged 4 and over in the Basque Autonomous Community (BAC) (data 1991–2021)
- The Arrue Report (analysing the use of the Basque language among school pupils)
- The Census
- Measuring language use on the street via street surveys

Please note that variations in the locations where the data is collected mean that there are differences in the degrees of the linguistic picture collected across the different areas where Basque is spoken, with the most extensive range of data for the Basque Autonomous Community (BAC).

4.4.2 A notable element of these arrangements is that there is more emphasis here on combining data from quantitative surveys with data collected through observation as a result of the use of street surveys that are considered a highly important source and the Arrue project examining the use of Basque in schools.

4.4.3 In terms of territorial levels of data collection, there is an emphasis on data collection to obtain a macro level understanding but also on the value of data at a more micro or urban level (municipal level) to feed into the indicator system. Collection of data at a more micro level is considered to be one of the main ways that the data sources have increased over time as more local data enables a better understanding of day-to-day situations.

4.4.4 Consideration has been given here again to the potential of collecting data regarding individuals' linguistic habits in day-to-day tasks such as using cash machines, paying for parking, or online. The need to collaborate with different organisations to make the most of big data was noted.

4.4.5 As part of implementing language indicator system, specific work groups are held, and this highlights the type of work areas that are considered important in this case and that require more detailed attention:

- University – the extent to which degree and postgraduate level studies are offered by language;
- Sports and leisure – including sports at school, being able to speak Basque among sports coaches;
- Media – the use of the media in Basque on a day-to-day basis;
- The digital world – websites in Basque, the use of software in Basque etc.;
- By highlighting these work streams, attention is drawn to the type of areas where specific data is available to feed into their work.

4.5 General points

In terms of data collection needs and processes, a number of relevant general points were made in the various discussions:

- If a system of language indicators is established, arrangements need to be in place to ensure that there is a clear relationship between the indicators and measurable data that is regularly available over a period of time.
- Surveys and questions that produce specific data need to be asked regularly so that trends can be assessed over time.
- There is a need to standardise the type of questions asked for a specific language. It is important to avoid a situation where the use of different questions

by various organisations and agencies do not exacerbate difficulties in comparing data across different sources and over time.

- The particular value of data to understand more micro and local contexts needs to be noted if we want to understand the progress and problems associated with the language in these contexts and allow for purposeful language planning.
- An element of pragmatism is needed for these processes in terms of refining and standardising data that is already being collected, and to consider the resources implications of seeking to collect more data.

5.0 Understanding and analysing language indicators

5.1 Technology and data mapping

5.1.1 Technology and data mapping are utilised in many policy areas and are also widely used in the private sector. The main advantage of these technologies is to create alternative representations of a social or economic phenomenon when different variables are combined. This produces a deeper understanding of the relationship between variables, enabling those who have produced or commissioned the representations to plan more purposefully and intentionally for the future. A simple example is the way such technologies are used in decisions on the location of new supermarket stores based on a number of factors, including: the number and density of population; the economic and employment profile of the population; the price and availability of suitable land for site development; the proximity of other supermarkets; transport systems.

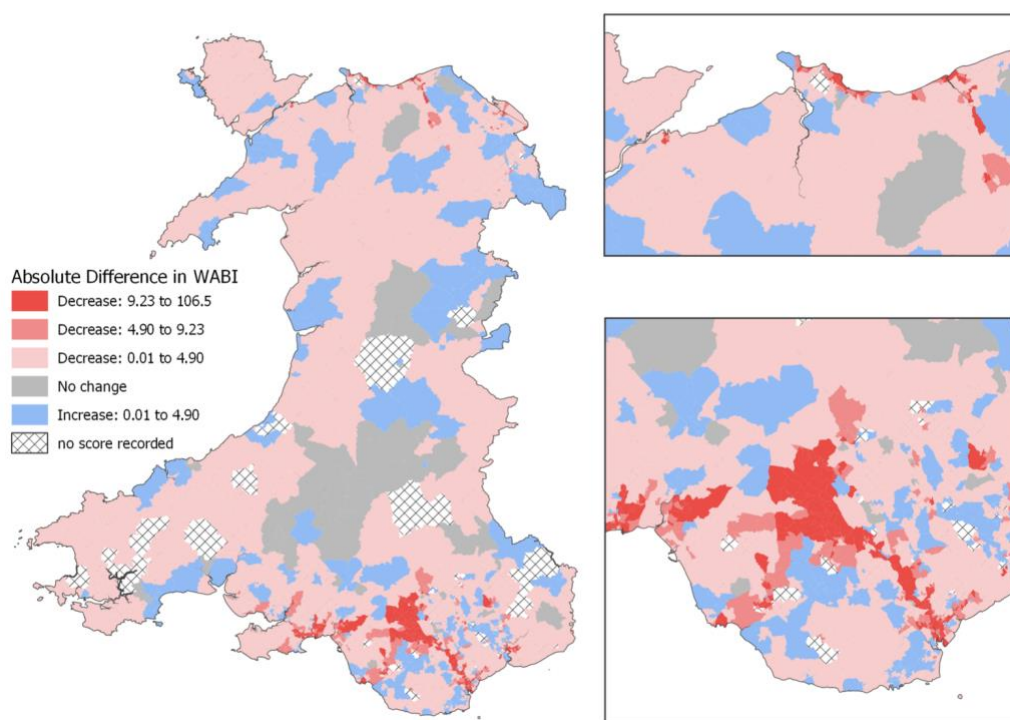
5.1.2 Overall, in Wales, the use of technology and mapping data in order to understand and interpret Welsh linguistic vitality is currently limited (Jones and Lewis 2019). The census data is the main source for creating fairly simple maps of linguistic ability, and these maps mainly focus on representing national patterns of percentages of Welsh speakers. Other maps of the geography of the Welsh language are produced by agencies such as the Welsh Language Commissioner. We understand that these methods will be used by the Bro Project. There is value in further investigating the potential of the National Statistics Service's digital data visualisation work in developing interactive maps in relation to the 2021 Census and whether that could be developed further for the purposes of understanding the Welsh language. So far, in general, there is little imaginative use of technologies and mapping data in order to understand and communicate Welsh linguistic vitality.

5.1.3 In discussions with stakeholders from Wales, there was consensus regarding the dire need for alternative ways of mapping the Welsh language, with a leading role for using technology. There was a sense of great potential for instance in using Geographical Information Systems (GIS) to gain a more dynamic understanding of the linguistic situation and to map the vitality of the Welsh language. The value of GIS as a method and as a technology is that it enables a number of social, economic and cultural variables to be mapped at the same time, enabling

researchers and others to evaluate the relationship between those variables and the language.

5.1.4 The figure below is a recent example of the use of GIS in Wales. The map represents the accessibility of bus services on a local scale in Wales and, as a result, the map combines information about bus routes with the proximity of individuals to bus stops. The map, therefore, illustrates current inequalities in the ability of individuals to gain access to use public transport, together with possible locations for future investment in public transport.

Figure 2. Change in WABI score, weekday services, 2019 – 2021. The upper map shows the absolute score differences, the lower map the percentage change.



5.1.5 The above example illustrates the potential of using GIS for interpreting Welsh linguistic vitality. GIS could be used as a way of combining different language variables in order to examine the spatial relationship between them. More importantly, GIS provides an opportunity to explore the spatial relationship between different linguistic variables and other social or economic variables. This can portray the current vitality or vulnerability of the language in a specific location, as well as being an opportunity to identify where proposed investments could potentially increase vitality in the future.

5.1.6 In summary then, the use of a GIS could be of assistance in:

- developing new insights, that could lead to more specific and targeted policies and interventions that would ultimately be more effective.

- developing new representations of the vitality of the language, helping to change the discourse from one of decline to a much more complex one of relative vitality and relative vulnerability.

5.1.7 It is anticipated that there are two ways of using GIS in the context of the Welsh language. Firstly, and quite simply, GIS could be used as a way of producing maps or representations of various variables linked to the vitality of the Welsh language, for example for a specific area. These different representations can be used as the subject of a qualitative discussion that would be a means to evaluate the vitality of the Welsh language.

5.1.8 Secondly, and in a way that reflects the true potential of GIS as a method, GIS could be used as a way of creating more composite and more statistical measures of the vitality of the Welsh language. In order to use GIS in this way, it would be necessary to complete some preliminary work, specifically developing a detailed understanding of the relationship between different variables or indicators, including weighting different elements and identifying causality between those variables (see part 2 of this report and comments 2.2.13 in particular). Although there are challenges associated with this option, there would also be some benefits. If the model forming the basis of the composite measure is accepted, this method could offer greater certainty when considering the vitality of the Welsh language, e.g. when determining the impact of a proposed development on the vitality of the Welsh language in an area or when assessing changes in the vitality of the Welsh language from one period to another and from one area to another. This method could also potentially model future vitality.

5.1.9 As part of the research carried out in the preparation of this report, the need to explore the potential of using artificial intelligence was identified as a way of generating new insights into the vitality of language, including the Welsh language. For example, artificial intelligence would enable the analysis of various databases, together with the potential to combine different databases. It was noted that further work is needed to explore the potential of this new technology in this context.

5.2 Good practice and useful examples internationally

5.2.1 There is some evidence of the use of technology and data mapping in other countries, and particularly in the Basque Country. GIS is used in the Basque Country as part of the EAS system. The EAS system allows users to produce maps of various variables on demand. Nevertheless, it is not currently a system that enables users to combine different variables together, thus reducing the ability to analyse the language situation.

5.2.2 Another example is language planners in the Basque Country collaborating with architects to study the relationship between the socio-linguistic situation and other metrics, especially the evolution of towns, e.g. building major roads, shopping centres, housing centres. As part of this work, two towns were compared - one that

had experienced an increase in the use of the language and another where use was decreasing - in order to interpret whether there was a possible relationship between the form of the town and linguistic changes. The expectation is that the results of this study will be published at the end of 2024.

5.2.3 Generally, all stakeholders who were consulted as part of the project agreed that further consideration should be given to the potential of technology and data mapping in providing effective ways of illustrating linguistic vitality. Research to this effect was underway in Catalonia.

5.2.4 In terms of developing capacity, there is considerable potential to share good practice in this area, especially bearing in mind the specialist skills needed to develop bespoke language models for GIS.

Bibliography

Baker, C. 2011. *Foundations of Bilingual Education and Bilingualism*. Clevedon: Multilingual Matters.

Baker, C., Wright, W. 2021. *Foundations of Bilingual Education and Bilingualism*. Clevedon: Multilingual Matters.

Brugera Soler, C., Ferret Baig, A., Torrijos López, A. 2022. 'El Sistema d'Indicadors Lingüístics: diagnòstic i avaluació en política lingüística' *Treballs de Sociolingüística Catalana*, num 32, pp. 247–260. DOI: 10.2436/20.2504.01.197

Cunliffe D. 2021. Minority languages in the age of networked individualism: from social networks to digital breathing spaces. In: Lewis H, McLeod W (eds) *Language Revitalisation and Social Transformation*. London: Palgrave Macmillan, 67–97.

Edwards, J. 1992. Sociopolitical Aspects of Language Maintenance and Loss: Towards a Typology of Minority Language Situations. In *Maintenance and Loss of Minority Languages*, ed. W. Fase, K. Jaspaert and S. Kroon, 37–54. Amsterdam: John Benjamins.

Fishman, J. A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon: Multilingual Matters.

Fishman, J. A., ed. 2001. *Can Threatened Languages be Saved? Reversing Language Shift, Revisited: A 21st Century Perspective*. Clevedon: Multilingual Matters.

Gazzola, M., Iannàccaro, G. 2024. 'Indicators in language policy and planning' in M. Gazzola, F. Grin, L. Cardinal and K. Heugh ed. *The Routledge Handbook of Language Policy and Planning*. London: Routledge.

A paper forming the basis of this chapter produced in 2021: https://www.ulster.ac.uk/_data/assets/pdf_file/0019/832060/REAL21-2.pdf

Giles, H., Bourhis, R. Y., and Taylor, D. M. 1977. Towards a theory of language in ethnic group relations. In *Language, Ethnicity and Intergroup Relations*, ed. H. Giles, 307–348. London: Academic Press.

Grenoble, L. 2013. Language revitalisation. In *The Oxford Handbook of Sociolinguistics*, eds. R. Bayley, R. Cameron, C. Lucas, 792–811. Oxford: Oxford University Press.

Grenoble, L. and Whaley, L. 2006. *Saving Languages: An Introduction to Language Revitalization*. Cambridge: Cambridge University Press.

Harbert, W. 2011. Endangered languages and economic development. In *The Cambridge Handbook of Endangered Languages*, eds. P. K. Austin a J. Sallabank, 403–422. Cambridge: Cambridge University Press.

Hornberger, N. 2010. Language shift and language revitalisation. In *The Oxford Handbook of Applied Linguistics*, ed. R. B. Kaplan, 412–420. Oxford: Oxford University Press.

Huss, L. 2017. Researching language loss and revitalisation. In *Research Methods in Language and Education*, ed. K. A. King and N. H. Hornberger, 99–111. New York: Springer.

Jones, R. and Lewis, H. 2019. *New Geographies of Language: Language, Culture and Politics in Wales*. London: Palgrave Macmillan.

I

King, K. A. 2001. *Language Revitalization Processes and Prospects*. Clevedon: Multilingual Matters.

Lewis, M.P. and Simons, G. 2010. Assessing endangerment: Expanding Fishman's GIDS, *Revue Roumaine de Linguistique* 55: 103–120.

Lewis, M.P. and Simons, G. 2016. *Sustaining Language Use: Perspectives on Community-based Language Development*. Dallas TX: SIL International.

Lewis, H. and Royles, E. 2018. Language revitalisation and social transformation: evaluating the policy frameworks of sub-state governments, *Policy and Politics* 46(3): 503–529.

Lewis, H. and McLeod, W. (eds) 2021. *Language Revitalisation and Social Transformation*. London: Palgrave Macmillan.

Pérez Báez, G., Vogel, R. and Koller, E. O. 2018. Comparative analysis in language revitalisation practices. In *The Oxford Handbook of Endangered Languages*, eds. K. L. Rehg and L. Campbell, 466–489. Oxford: Oxford University Press.

Pérez Báez, G., Vogel, R. and Patolo, U. 2019. Global survey of revitalization efforts: A mixed methods approach to understanding language revitalization practices, *Language, Documentation and Conservation* 13: 446–513.

Solé Camardons, J. 2003. 'El Sistema d'Indicadors Lingüístics (SIL): finalitat i característiques generals,' *Revista de Sociolingüística* Estiu 2003, https://www.gencat.cat/llengua/noves/noves/hm03estiu/sole1_3.htm

Spolsky, B. 1996. Conditions for language revitalisation: A comparison of the cases of Hebrew and Maori. In *Language and the State: Revitalisation and Revival in Israel and Eire*, ed. S. Wright, 177–201. Clevedon: Multilingual Matters.

Tsunoda, T. 2005. *Language Endangerment and Language Revitalization*. Berlin: Walter de Gruyter.

UNESCO. 2003. *Language Vitality and Endangerment: Ad Hoc Expert Group on Endangered Languages*. Document adopted by the International Expert Meeting on UNESCO Programme Safeguarding Endangered Languages, Paris, March 10–12.

Welsh Government. 2022. *Cymraeg 2050: A million Welsh speakers Research and Evaluation Framework*. Cardiff: Welsh Government. Available from: <https://www.gov.wales/sites/default/files/statistics-and-research/2022-11/cymraeg-2050-a-million-welsh-speakers-research-and-evaluation-framework.pdf>

Williams, C.H. 2022. 'Spatial representation and sociolinguistic synergies' *Sociolinguistica* 36,1–2, 231–245 <https://doi.org/10.1515/soci-2022-0016>

Appendix 1: Information on those consulted during the course of this project (in alphabetical order)

| | |
|---------------------------|-----------------------------|
| Bonner, Elen | Bangor University |
| Cunliffe, Daniel | University of South Wales |
| Evas, Jeremy | Welsh Government |
| Gazzola, Michele | Ulster University |
| Hodges, Rhian | Bangor University |
| Iorwerth, Hywel | Welsh Language Commissioner |
| James, Eleri | Welsh Language Commissioner |
| Jenkins, Dave | Aberystwyth University |
| Jones, Eleri | Welsh Government |
| Jones, Kathryn | Cwmni Iaith Cyf |
| Jones, Meirion Prys | Lingua Ni |
| Jones, Patience | Welsh Government |
| Larrea, Immanol | Soziolinguistika Klusterra |
| O'Rourke, Bernadette | Glasgow University |
| Parry, Martin | Welsh Government |
| Prys, Cynog | Bangor University |
| Sendra Rovira, Montserrat | Generalitat de Catalunya |
| Torrijos Lopez, Anna | Generalitat de Catalunya |
| Walsh, John | Foras na Gaeilge |
| Williams, Colin | Cardiff University |
| Wyn, Owain | Burum |