

AURKEZPENA

XXI. mendean aurrera goaz, eta ingurune digitala gure bizitzaren parte da, hori ez du inork zalantzan jartzen. Adimen artifizialak eta, bereziki, hizkuntza-teknologiak gure egunerokoan sartuak ditugu, eta euskara ez da salbuespena. Horregatik antolatu nahi izan du Euskaltzaindiak, bere mendeurrenaren egitarau akademikoaren barruan, «Ingurune digitala, hizkuntzen estandarizazioa eta euskara» izeneko jardunaldia. Eta hori galdera baten bidez egin du: nola lagun dezake ingurune digitalak euskararen estandarizazioa bizkortzen? Horretarako, beste hizkuntzetako esperientziak, Euskaltzaindiarenak berarenak eta euskal erakundeek ezagutu nahi izan ditugu, guztiok elkarrekin euskararen estandarizazio horren formulazioan, ebaluzioan eta zabalkundearen aurrera egiteko eta XXI. mendeko euskal gizarteak dituen euskarari buruzko premiei erantzun egokia emateko. Azken batean, Akademiak ezinbestekoa duelako inguruan gertatzen dena ezagutzea, eta hizkuntzaren estandarizazioan egindakoa neurtzea.

Lau atal nagusitan egituratu da liburua.

Lehenengoan XXI. mendeko euskararen estandarizazioaren nondik norakoak aztertzen dira: baldintzak, ezaugarriak eta ingurune digitala. **Andoni Sagarna** euskaltzain oso eta Euskaltzaindikoko IKT batzordeko buruak euskararen estandarizazioan baliabide digitalak nola erabili izan diren eta etorkizunera begira zer erronka eta aukera aurreikusten diren aurkezten du bere artikuluan. Euskararen estandarizazioa 1968an hasi zenetik igaro diren 50 urte baino gehiagoan, baliabide digitalen erabilera asko aldatu da. Artikuluan bi prozesu horiek, euskararen estandarizazioak eta baliabide digitalen erabilerak, izan duten bilakaera eta eragin-trukea deskribatzen ditu. Aldaketa horien barruan bereziki nabarmentzen da corpusen erabilera hizkuntzaren ikerketan.

Bigarren atalean akademien ikuspegitik heltzen zaio ingurune digitalak hizkuntzaren estandarizazioan duen lekuari, eta akademien zereginaz dihardute hainbat adituk. **Mathieu Constant** Lorraineko Unibertsitateko ikertzaileak hitz anitzeko unitate lexikalak corpusetan oinarrituz identifikatzeko hainbat prozeduraren berri ematen du, frantseserako. Identifikazio automatikorako hiru metodo aurkezten ditu, eta, azkenik, corpus anotatu baten bidezko zein baliabide lexikalen bidezko ikasketaz baliatuz unitate konplexu horiek identifikatzeko metodo bana eta bien konbinazio bidezkoa aurkezten ditu. **Benazet Dazéas** Lo Congrès Permanent de la Lengua Occitanako zuzendaria eta 2012az geroztik *locongres.org* atariko arduraduna da hurrengo artikulua egilea. Hainbat eduki biltzen ditu atariak, nagusia *dicod'Òc* 'multihiztegia' deitzen dutena bada ere. Haatik, edukiak egituratu gabe eta elkarrekin lotu gabe daudela oharturik, multiplikazio bat sortzen ari dira lexikografia, flexioa, toponimia edo corpusak elkarrekin konektatu eta kontsulta bakarrean eskaini ahal izateko. Ondoren, **Manuel González** Real Academia Galegako osoko kide eta Santiagoko Unibertsitateko filologia erromanikoko katedradunak lexikoaren estandarizazioaz eta Galiziako Akademiaren hiztegiak dihardu bere artikuluan. Hiztegi horren idazketan, berrikuspenean eta zabalkundearen baliabide digitalen erabilerak duen garrantzia agertzen du. **Joaquim Rafel** Institut d'Estudis Catalaneko Secció Filològicako osoko kide eta Bartzelonako Unibertsitateko katedradun emerituak deskribatzen du, jarraian, hizkuntzalaritzak eta, arlo horretan, informatikaren erabilerak azken urteetan izandako aurrerapenek lexikografiari ekarri dioten egoera-aldaketa, batez ere paperezko hiztegietatik informatizatuetarako jauziak dakarrena. Hiztegien kontzepzioa bera ere nola ari den aldatzen azpimarratzen du, bai sintaxiarekiko harremanaren aldetik eta baita kombinatoria lexikalarenetik ere. Lexikografia katalanak xx. mendean izan zuen bilakaeraren balantzea ere egiten du bere artikuluan. **Guillermo Rojo** Espainiako Errege Akademiako osoko kide eta Santiago de Compostelako Unibertsitateko katedradun emerituarena da hurrengo artikulua. RAEk azken 25 urtean bere lanak informatizatzeko egindako ahaleginen fruitu oparoak aurkezten ditu Rojok, besteak beste corpus-hizkuntzalaritzan oinarritutako lanak biltzen eta harremanetan jartzen dituen *Enclave RAE*ren garrantzia azpimarratzen du. Espainiako Errege Akademiak egin dituen beste digitalizazio-lan batzuen berri ere ematen du, baita bere zuzendaritzapeko corpusen ezaugarriak xehetasunez azaltzen ere, akademiak sortu dituen kontsulta-zerbitzuak ahaztu gabe. Akademien ikuspegiari eskainitako bigarren atal honetako azken artikulua **Hans Van de Velde** Fryske Akademyko kide eta Utrecht-eko Unibertsitateko irakaslearena da. Frisieraren normalizazioaz gain, garapen digitalari buruzko gogoetaren berri eman eta hizkuntzaren ikusgarritasunaren garrantzia aipatu ondoren, frisierarako garatu diren tresna nagusiak aurkezten ditu artikuluan. Bukatzeko, aro digitaleari bereziki eta hobe-

to landu beharreko ikerketa eta garapena zein hizkuntza-politikari buruzko ideia batzuk aurkezten ditu.

Hirugarren atala estandarizazioa gizarteratzerakoan ingurune digitalak duen lekuari eskainia dago, eta hainbat esperientziaren berri ematen da. **Iolanda Fernández** Real Academia Galegako komunikazio-arduradunak akademian azken urteotan egindako bideari buruzko ideia interesgarriak dakarzkigu bere lanean. Iraultza digitalak RAGen estandar arauemailearen zabalkundeari aukera berriak ekarri dizkio. 2012an hiztegia sarean jarri zuenetik kontsultek etengabe egin dute gora, baina akademiaren apustua harago doa, eta eduki ludikoak eta didaktikoak ere eskaintzen ditu. Ez hori bakarrik: aukera horiek baliatu dira hizkuntzaren estandarren berri emateko, baina bereziki hiztunekin, ikasleekin eta profesionalekin harremana estutzeko, haien beharrak entzun eta erantzunak doitzeko batetik, eta galegoarekiko jarrera hobetzeko bestetik, hizkuntza minorizatua dela eta diglosia-egoeran dagoela ezin baita ahaztu. **Vincent Rivière** Lo Congrès Permanent de la Lengua Occitanako hizkuntza-proiektuen arduradunak gaskoiaren aditzak jokatzekeo tresna bat aurkezten du erabiltzaile-talde zabalarentzat. Berrikuntza handia dakar horrek, orain arte ez baitzegoen halako tresnarik gaskoirako. **Gilles Souvay** Analyse et Traitement Informatique de la Langue Française (ATILF) ikerketa-laboregiko kideak hiztegi elektronikoei buruz dihardu bere artikuluan, ATILFek landu baitu hasieratik *Trésor de la Langue Française*. Eta, hortik abiatuta, egun lantzen dituzten hainbat baliabide lexikografiko aurkezten ditu: hiztegiak (egungoak, historikoak, etimologikoak), testu-corpusak eta hizkuntzen tratamendu automatikorako bestelako tresnak eta baliabideak.

Eta, bukatzeko, laugarren atalean euskararen estandarizaziorako ingurune digitalaren baliabide linguistikoak eta sozialak ditugu aztergai. Euskarak ingurune digitalean duen osasuna aztertzekeo egin duen ariketa baten emaitzak aurkezten dizkigu **Ainhoa Arregi** UZEIko lexikografia saileko teknikari eta euskaltzain urgazleak, euskararen adierazleak oinarri hartuta. Araugintzaren garrantziarekin batera, hura azkar eta egoki zabaltzekeo garrantziaz dihardu, eta profesionalen arduraz. **Josu Aztiria** Elhuyar Fundazioko Hizkuntza eta Teknologia unitateko zuzendariak adimen artifizialaren aroan euskarak dituen aukera-leihoak aipatzen ditu, estrategia sinbolikotik ikasketa automatikora egin diren urratsak erakutsita. Etorkizun hurbilean landu beharreko hainbat interbentzio-esparru ere proposatzen ditu. Ondoren EHUko IXA taldeko kide diren **Olatz Arregi**, **Iñaki Alegria** eta **Xabier Artola** irakasleek estandarizazioaren eta hizkuntza-teknologiaren inguruko gogoeta egiten dute, bidean ikasitakoaz eta plangintza baten beharraz. Ordenagailuak, digitalizazioa, Internet, liburu elektronikoak, telefono adimendunak eta sare sozialak, besteak beste, gure etxeetan eta enpresetan sartuz joan dira, eta horiekin batera baita hizkuntza-teknologiak ere.

Euskara goiz hasi zen arlo teknologikora egokitzen, eta teknologia ere hasi zen euskara kontuan hartzen, baina baliabide handiagoko hizkuntzekin alderatuta egoera «osatugabea» dago oraindik. Hala, bada, etorkizunerako dauden erronkak oso handiak dira. Halako plangintza baten nondik norakoak eta balizko lerro nagusiak proposatzen dituzte artikuluan. Atal honetako azken artikulua-ren egilea **Jean-Baptiste Battitu Coyos** euskaltzain oso eta Jagon sailburua da. Coyosek Euskaltzaindiaren *Euskara eskuz esku* (EEE) baliabide berria aurkezten du artikulua honetan. EEEren helburua euskara Iparraldean hobeki erabiltzeko laguntza eskaintzea da, bertako berezitasunak bistaratu eta, horrekin batera, Euskaltzaindiaren arauak modu sinplifikatuan aurkeztea, horretarako berariazko webgunea prestatuta.

Esperientzia hauek guztiak ezagututa, gogoetarako garaia dator, hasierako galderari erantzuna emateko tresnak eskura ditugu eta. Hori izango da etorkizun hurbilerako erronka.

PRESENTACIÓN

Estamos avanzando en el siglo XXI, y el entorno digital es ya parte de nuestras vidas; nadie lo pone en duda. La inteligencia artificial y particularmente las tecnologías de la lengua se nos han introducido en el día a día, y el euskera no es una excepción en esto. Ese es el motivo por el cual Euskaltzaindia ha querido organizar, dentro del programa académico conmemorativo del centenario de su creación, la jornada «El entorno digital, la estandarización de la lengua y el euskera». Y la ha planteado mediante una pregunta: ¿cómo puede ayudar el entorno digital a dinamizar la estandarización del euskera? Con ese fin, hemos querido conocer las experiencias de otras lenguas, las de la propia Euskaltzaindia y las de las instituciones vascas, para avanzar conjuntamente en la formulación, evaluación y difusión de esa estandarización del euskera y dar una respuesta correcta a las necesidades que la sociedad vasca tiene planteadas con relación a la lengua vasca. Porque, a fin de cuentas, a la academia le es imprescindible conocer lo que ocurre en su entorno, y medir la labor realizada en la estandarización de la lengua.

El libro se ha estructurado en cuatro partes. En la primera se analiza el devenir de la estandarización del euskera en el siglo XXI: las condiciones, las características y el entorno digital. **Andoni Sagarna**, académico de número y responsable de la comisión TIC de Euskaltzaindia describe en su artículo el uso que se ha hecho de los recursos digitales, y plantea los retos y las oportunidades con vistas al futuro. En los más de 50 años transcurridos desde el inicio de la estandarización del euskera en 1968, el uso de los recursos digitales ha evolucionado mucho. En su artículo describe la evolución y las interacciones de la estandarización del euskera con los recursos digitales. En ese desarrollo se destaca el uso de corpus en la investigación de la lengua.

En la segunda parte varios expertos analizan el lugar que ocupa el entorno digital en la estandarización de la lengua y las tareas de las academias en ese ámbito. **Mathieu Constant**, investigador de la Universidad de Lorena, ha presentado diversos procedimientos de identificación de unidades léxicas complejas

basándose en corpus del francés. Da cuenta de tres métodos de identificación automática, y, finalmente, presenta dos métodos para identificar esas unidades complejas, uno por medio de un corpus anotado, otro de aprendizaje mediante recursos léxicos y el consistente en la combinación de ambos. **Benazet Dazéas**, director de Lo Congrès Permanent de la Lengua Occitana y desde 2012 responsable del portal locongres.org, es el autor del siguiente artículo. El portal reúne varios contenidos, entre los que destaca el ‘multidiccionario’ *dicod’Òc*. Habiendo observado que los contenidos no estaban estructurados ni interconectados, están creando una multiaplicación para interrelacionar la lexicografía, la flexión, la toponimia y los corpus, y poder ofrecer una consulta unificada. A continuación, **Manuel González**, académico de número de la Real Academia Galega y catedrático de Filología Románica de la Universidad de Santiago de Compostela, presenta en su artículo la estandarización del léxico y la labor de la Academia Gallega. Destaca la importancia que tiene el uso de los recursos digitales en la redacción, revisión y difusión de su diccionario. **Joaquim Rafel**, miembro de número de la Secció Filològica del Institut d’Estudis Catalans y catedrático emérito de la Universidad de Barcelona describe, a continuación, el cambio al que han dado lugar en el campo de la lexicografía los avances que se han producido en la lingüística y en el uso de la informática, sobre todo el que se ha dado en el salto de los diccionarios en formato papel a los informatizados. Subraya en particular cómo está cambiando la propia concepción de los diccionarios, tanto desde el punto de vista de su relación con la sintaxis como de la combinatoria léxica. En su artículo también hace balance de la evolución experimentada por la lexicografía catalana en el siglo xx. El siguiente artículo es de **Guillermo Rojo**, académico de número de la Real Academia Española y catedrático emérito de la Universidad de Santiago de Compostela. Rojo presenta los abundantes frutos que la RAE ha obtenido en los últimos 25 años, como resultado de los esfuerzos realizados en la informatización de sus trabajos. Subraya la importancia de *Enclave*, el sistema que reúne y relaciona los trabajos de la RAE en el campo de la lingüística de corpus. Da cuenta también de otros trabajos de digitalización de la Real Academia Española y describe en detalle las características de los corpus que dirige, así como los servicios de consulta creados por la academia. El último artículo de esta parte dedicada a la visión de las academias es de **Hans Van de Velde**, miembro de la Fryske Akademy y profesor de la Universidad de Utrecht. Después de presentar la normalización de la lengua frisía, ofrece su visión en torno al desarrollo digital, subraya la importancia de la visibilización de la lengua y presenta las principales herramientas que se han desarrollado para la lengua frisía. Finalmente reflexiona sobre aspectos que es necesario mejorar en la investigación y el desarrollo de la era digital así como en la política lingüística.

La tercera parte está dedicada al lugar que ocupa el entorno digital en la difusión social de la estandarización, y se muestran diversas experiencias de ese ámbito. **Iolanda Fernández**, responsable de comunicación de la Real Academia Galega, da a conocer en su trabajo varias ideas interesantes acerca del camino recorrido a lo largo de estos últimos años por la academia. La revolución digital ha aportado nuevas oportunidades para la difusión de la labor estandarizadora de la RAG. Desde que publicó el diccionario en la red en 2012, las consultas han crecido constantemente, pero la apuesta va más allá, porque ofrece también contenidos lúdicos y didácticos. Además se han utilizado esos recursos para informar de la lengua estándar, y, sobre todo, para estrechar las relaciones con los hablantes, los estudiantes y los profesionales, para conocer sus necesidades y ajustar las respuestas a aquellas por un lado, y, por otro lado, para mejorar las actitudes respecto del gallego, ya que no se puede olvidar que se trata de una lengua minorizada y se encuentra en situación de diglosia. **Vincent Rivière**, responsable de los proyectos lingüísticos de Lo Congrès Permanent de la Lenga Occitana, presenta una herramienta para la conjugación de los verbos del gascón para un amplio grupo de usuarios. Supone una innovación muy importante, puesto que hasta ahora no existía ese tipo de herramientas para el gascón. **Gilles Souvay**, miembro del laboratorio de investigación Analyse et Traitement Informatique de la Langue Française (ATILF), trata en su artículo de los diccionarios electrónicos que han desarrollado. El ATILF ha elaborado la versión electrónica del *Trésor de la Langue Française*. A partir de ahí han desarrollado diversos recursos lexicográficos: diccionarios (modernos, históricos, etimológicos), corpus textuales y otros recursos y herramientas para el tratamiento automático de la lengua.

Y, finalmente, la cuarta parte trata de los recursos lingüísticos y sociales del entorno digital para la estandarización del euskera. **Ainhoa Arregi**, lexicógrafa de UZEI y académica correspondiente de Euskaltzaindia, nos muestra los resultados de un ejercicio que ha realizado para estudiar la salud de la que goza el euskera en el entorno digital, tomando como punto de partida algunos indicadores relacionados con el euskera. Concede importancia no solo a la normativización sino también a la difusión rápida y adecuada de la misma, y subraya la responsabilidad de los profesionales a este respecto. **Josu Aztiria**, director de la Unidad de Lengua y Tecnología de la Fundación Elhuyar, expone las ventanas de oportunidad que el euskera tiene en la era de la inteligencia artificial, y da cuenta de los avances realizados desde la estrategia simbólica al aprendizaje automático. También propone diversos campos de intervención en los que hay que trabajar. A continuación, **Olatz Arregi, Iñaki Alegria y Xabier Artola**, profesores de la UPV y miembros del grupo IXA reflexionan sobre la estandarización y las tecnologías lingüísticas, lo aprendido en el camino recorrido y la necesidad de

planificación. Los ordenadores, la digitalización, Internet, los libros electrónicos, los teléfonos inteligentes y las redes sociales entre otras cosas han entrado en nuestros hogares y empresas, y junto con esas cosas también las tecnologías de la lengua. El euskera se empezó a adaptar al campo de la tecnología, y la tecnología empezó a tener en cuenta al euskera tempranamente, pero si comparamos su situación con la de las lenguas que gozan de mayores recursos, se encuentra todavía en una situación de «incompletitud». Así pues, los retos con vistas al futuro son muy grandes. En su artículo proponen también la orientación y las líneas maestras de una planificación. El autor del último artículo de esta parte es **Jean-Baptiste Battittu Coyos**, académico de número de Euskaltzaindia y responsable de la Sección Tutelar de la academia. Coyos presenta en su artículo el nuevo recurso de Euskaltzaindia denominado *Euskara eskuz esku* (EEE) El objetivo del EEE es ofrecer una ayuda para un mejor uso del euskera estándar en Euskal Herria continental, haciendo visibles las particularidades de aquella zona y, junto con eso, presentar las normas de Euskaltzaindia de una manera simplificada, mediante un sitio web desarrollado específicamente con esa finalidad.

Conocidas todas esas experiencias, ha llegado la hora de la reflexión, ya que disponemos de las herramientas para la búsqueda de la respuesta a la pregunta del principio. Ese va a ser el reto del futuro próximo.

PRÉSENTATION

Alors que nous avançons dans ce xxi^e siècle, l'environnement numérique fait déjà partie de nos vies, nul ne peut en douter. L'intelligence artificielle et plus particulièrement les technologies de la langue se sont insinuées en nous jour après jour et la langue basque ne fait pas exception. C'est la raison pour laquelle Euskaltzaindia a souhaité organiser, dans le cadre des manifestations commémoratives du centenaire de sa création, un colloque intitulé «L'environnement numérique, la standardisation de la langue et la langue basque». Ce colloque a été organisé autour d'une question : comment l'environnement numérique peut-il aider à dynamiser la standardisation du basque ? Nous avons pour cela souhaité connaître les expériences des autres langues, ainsi que celle d'Euskaltzaindia et des institutions basques, pour avancer conjointement dans la formulation, l'évaluation et la diffusion de la standardisation du basque et apporter une réponse adéquate aux besoins relatifs au basque exprimés par la société basque. Car, en fin de compte, l'Académie a absolument besoin de connaître ce qui se passe autour d'elle et de mesurer le travail réalisé en matière de standardisation de la langue.

Ce livre se divise en quatre parties.

Dans la première partie est analysé le devenir de la standardisation du basque au XXI^e siècle : les conditions, les caractéristiques et l'environnement numérique. **Andoni Sagarna**, académicien titulaire responsable de la commission TIC d'Euskaltzaindia (Technologies de l'Information et de la Communication), décrit dans son article l'usage qui a été fait des ressources numériques, il évoque les enjeux et les opportunités pour l'avenir. Au cours des 50 années passées depuis le début de la standardisation du basque en 1968, l'usage des ressources numériques a beaucoup évolué. Il décrit dans son article l'évolution et les interactions de la standardisation du basque avec les ressources numériques. On constate l'utilisation du corpus dans la recherche linguistique.

Dans la seconde partie plusieurs experts analysent la place qu'occupe l'environnement numérique dans la standardisation de la langue et les travaux menés par les académies dans ce domaine. **Mathieu Constant**, chercheur à l'Université de Lorraine, présente divers procédés d'identification des unités lexicales complexes basés sur le corpus du français. Il présente trois méthodes d'identification automatique, ainsi que deux méthodes pour identifier ces unités complexes, l'une grâce à un corpus annoté et l'autre étant un apprentissage grâce à des ressources lexicales et consistant en la combinaison des deux. **Benazet Dazéas**, directeur de *Lo Congrès Permanent de la Lengua Occitana* et responsable depuis 2012 du portail *locongres.org*, est l'auteur de l'article suivant. Ce portail comprend plusieurs contenus parmi lesquels on peut remarquer le 'multidictionnaire' *dicod'Òc*. Ayant constaté que les contenus n'étaient ni structurés ni interconnectés, une multiapplication a été créée pour interconnecter la lexicographie, la flexion, la toponymie et les corpus, et pour permettre une recherche unifiée. **Manuel González**, académicien titulaire de l'Académie royale galicienne et professeur de philologie romane à l'Université de Saint-Jacques-de-Compostelle, présente quant à lui dans son article la standardisation du lexique, ainsi que le travail mené par l'Académie galicienne. Il souligne l'importance de l'usage des ressources numériques dans la rédaction, la révision et la diffusion de son dictionnaire. **Joaquim Rafel**, membre titulaire de la *Secció Filològica* de l'*Institut d'Estudis Catalans* et professeur émérite à l'Université de Barcelone, décrit ensuite les changements apportés en matière de lexicographie par les avancées de la linguistique et l'usage de l'informatique, surtout le passage des dictionnaires sur format papier vers les dictionnaires numériques. Il souligne en particulier le changement dans la conception elle-même des dictionnaires, tant du point de vue de sa relation avec la syntaxe que de sa combinaison lexicale. Il fait aussi dans son article un bilan de l'évolution qu'a connue la lexicographie catalane au cours du XX^e siècle. L'article suivant est de **Guillermo Rojo**, académicien titulaire de l'Académie royale espagnole et professeur émérite à l'Université de Saint-Jacques-de-Compostelle. G. Rojo présente les fruits abondants que l'Aca-

démie royale espagnole a récoltés au cours des 25 dernières années et qui sont les résultats des efforts réalisés en matière d'informatisation de ses travaux. Il souligne l'importance de *Enclave*, système qui rassemble et met en relation les travaux de l'Académie royale espagnole en matière de linguistique du corpus. Il évoque également d'autres travaux de numérisation de l'Académie royale espagnole et décrit en détail les caractéristiques des corpus qu'il dirige et les services de consultation créés par l'académie. Le dernier article de cette seconde partie consacrée à la vision des académies est de **Hans Van de Velde**, membre de la *Fryske Akademy* (académie frisonne) et professeur à l'Université d'Utrecht. Après avoir présenté la normalisation de la langue frisonne, il donne sa vision en matière de développement numérique, il souligne l'importance de la visibilité de la langue et présente les principaux outils qui ont été créés pour la langue frisonne. Il rend compte enfin de la réflexion qu'il a menée sur des aspects qu'il est nécessaire d'améliorer en matière de recherche et de développement de l'espace numérique mais aussi de politique linguistique.

La troisième partie est consacrée à la place occupée par l'environnement numérique dans la diffusion de la standardisation et où sont exposées plusieurs expériences dans ce domaine. **Iolanda Fernández**, chargée de communication à l'Académie royale galicienne, présente dans son texte plusieurs idées intéressantes sur le chemin parcouru par l'académie au cours de ces dernières années. La révolution numérique a apporté de nouvelles opportunités pour la diffusion du travail de standardisation mené par l'Académie royale galicienne. Depuis la publication en ligne du dictionnaire en 2012 les consultations n'ont cessé de croître, mais le pari va bien au-delà car il offre aussi des contenus ludiques et didactiques. Ces ressources ont également été utilisées pour apporter une information sur la langue standard et surtout pour renforcer les relations avec les locuteurs, les étudiants et les professionnels, pour connaître leurs besoins et ajuster les réponses d'une part, et d'autre part pour améliorer les comportements vis-à-vis du galicien, car il ne faut pas oublier qu'il s'agit d'une langue minorisée et qu'elle est dans une situation de diglossie. **Vincent Rivière**, responsable des projets linguistiques de *Lo Congrès Permanent de la Lengua Occitana*, présente un outil pour la conjugaison des verbes gascons pour un large groupe d'utilisateurs. Cela suppose une innovation très importante car jusque là ce type d'outil n'existait pas pour le gascon. **Gilles Souvay**, membre du laboratoire de recherche ATILF (Analyse et traitement informatique de la langue française), parle dans son article des dictionnaires électroniques qu'ils ont créés. L'ATILF a élaboré la version électronique du *Trésor de la langue française*. A partir de là ils ont créé diverses ressources lexicographiques : dictionnaires (modernes, historiques, étymologiques), corpus de textes et autres ressources et outils pour le traitement automatique de la langue.

Et enfin la quatrième partie parle des ressources linguistiques et sociales de l'environnement numérique pour la standardisation du basque. **Ainhoa Arregi**, lexicographe de UZEL et membre correspondante d'Euskaltzaindia, nous montre les résultats d'un exercice qu'elle a réalisé pour étudier la santé dont jouit la langue basque dans l'environnement numérique, prenant comme point de départ quelques indicateurs relatifs à la langue basque. Elle accorde de l'importance non seulement à la normativisation mais également à sa diffusion rapide et adéquate et souligne la responsabilité des professionnels en la matière. **Josu Aztiria**, directeur de l'Unité de langue et technologie de la fondation Elhuyar, présente les fenêtres d'opportunités dont bénéficie la langue basque dans le domaine de l'intelligence artificielle, et fait un compte rendu des avancées depuis la stratégie symbolique jusqu'à l'apprentissage automatique. Il propose également plusieurs champs d'intervention dans lesquels il faut œuvrer. Puis, **Olatz Arregi**, **Iñaki Alegria** et **Xabier Artola**, professeurs à l'Université du Pays Basque et membres du groupe IXA, mènent une réflexion sur la standardisation et les technologies linguistiques, ce qui a été appris sur le parcours et la nécessité de planification. Les ordinateurs, la numérisation, Internet, les livres électroniques, les téléphones intelligents et les réseaux sociaux entre autres choses, sont entrés dans nos foyers et nos entreprises et avec eux les technologies de la langue. La langue basque a commencé à s'adapter au domaine de la technologie et la technologie a commencé très tôt à prendre en compte la langue basque mais, si nous comparons sa situation à celle des langues qui bénéficient de meilleures ressources, elle est encore incomplète. Ainsi, les enjeux pour le futur sont très grands. Dans leur article ils proposent aussi l'orientation et les lignes directrices d'une planification. L'auteur du dernier article de cette quatrième partie est **Jean-Baptiste Battittu Coyos**, académicien titulaire responsable de la section de Sauvegarde de l'Académie. B. Coyos présente dans son article le nouvel outil d'Euskaltzaindia *Euskara eskuz esku* (EEE). L'objectif de EEE est d'offrir une assistance pour un meilleur usage du basque standard au Pays Basque Nord, mettant l'accent sur ses particularités, mais aussi de présenter les normes d'Euskaltzaindia d'une manière simplifiée, grâce à un site web créé spécifiquement dans ce but.

Toutes ces expériences étant maintenant connues, l'heure est venue de la réflexion puisque nous disposons des outils pour répondre à la question initiale. Cela va être l'enjeu dans un avenir proche.

PRESENTATION

We are well into the 21st Century and the digital environment is, without a doubt, part of our lives by now. Artificial intelligence, and language technol-

ogies in particular, have entered our everyday, and the Basque language is no exception here. This constituted the motive behind Euskaltzaindia's decision to organise the seminar «The Digital Environment: Language Standardization and Basque», as part of the academic program commemorating the centenary of its foundation. The seminar was conceived around a central question: how can the digital environment help revitalize the standardization of Basque? With this main objective in mind, we set out to learn about the experiences of other languages, that of Euskaltzaindia itself, and that of Basque institutions. The aim was to firstly move forward in the formulation, evaluation and circulation of Basque's standardization, to then be able to appropriately address the specific needs of Basque society in relation to the Basque language. For it is ultimately essential that the Academy know about what is going on in the world around it, and measure the work done in the standardization of language.

The book has been structured into four parts.

The first part analyzes the evolution of Basque's standardization in the 21st Century: the conditions and characteristics of this process and the role of the digital environment in it. **Andoni Sagarna**, member of Euskaltzaindia and head of the TIC (Communication and Information Technologies) Committee, describes in his article the different uses that have been made of digital resources and presents the challenges and opportunities for the future. It has been over 50 years since the standardization of Basque began in 1968, and the use of digital resources has evolved considerably since then. In his article, Sagarna offers a detailed account of the development of the standardization of Basque and its relationship with digital resources, highlighting the role of the use of corpus data in language research.

In the second part of the book, a number of experts analyze the role of the digital environment in language standardization and the work of language academies in this regard. **Mathieu Constant**, researcher at the University of Lorraine, has developed different procedures for the identification of complex lexical items drawing on the corpus data of the French language. He first describes three methods of automatic identification, and then introduces two different methods for identifying complex units: one via an annotated corpus, another learning method based on lexical resources, and one consisting of the combination of both. **Benazet Dazéas**, director of *Lo Congrès Permanent de la Lengua Occitana*, and in charge of the website *locongres.org* since 2012, authors the following article. One particular highlight among the website's contents is the 'multidictionary' *dicod-Ôc*. After observing that contents were neither structured nor interconnected, they are creating a multi-application in order to interrelate lexicography, inflection, toponymy, and corpus data, with the main objective of being able to offer the option of a unified query. In his article,

Manuel González, member of the Galician Royal Academy (RAG) and Professor in Romanic Philology at the University of Santiago de Compostela, discusses the standardization of lexicon and the work of the Galician Academy. He underscores the important role played by the use of digital resources in the writing, revision and dissemination of its dictionary. **Joaquim Rafel**, member of the *Secció Filològica* of the *Institut d'Estudis Catalans*, and Emeritus Professor at the University of Barcelona, describes in his article the changes in the field of lexicography resulting from the progress made in linguistics and the use of computer technology, especially the switch from paper dictionaries to digital ones. He lays a special emphasis on the changes that the dictionary as such, in the way it is understood and conceived, is undergoing, both from the perspective of its relationship to syntax, and from that of lexical combinatory. In his article he also takes stock of the evolution of Catalan lexicography in the 20th Century. The following article is by **Guillermo Rojo**, member of the Spanish Royal Academy (RAE) and Emeritus Professor at the University of Santiago de Compostela. In the article, Rojo presents the results of the RAE's efforts over the last 25 years to digitalize its work. He highlights the central role played by *Enclave*, a system that gathers and establishes links between the work that the RAE has developed in the field of corpus linguistics. Rojo also describes other digitalization projects by the RAE and gives us a detailed account of the specific characteristics of the corpus project that he is directing at the moment, and of the query services created by the RAE. The last article in this section focusing on the work of the national language academies is by **Hans Van de Velde**, member of the *Fryke Akademy* and lecturer at the University of Utrecht. After offering an account of the normalization of the Frisian language, he presents his understanding of digital development, highlighting the importance of language visibility and introducing the main tools developed for the Frisian language. He ends with a reflection on the different aspects that are in need of improvement, both in terms of research and development of the digital era, and in linguistic policy.

The third part of the book focusses on the role played by the digital environment in the social dissemination of standardization, and presents different case studies to explore this. In her text, **Iolanda Fernández**, head of communication at the Galician Royal Academy, provides a number of interesting insights about the path travelled by the Academy over the past few years. The digital revolution has provided new opportunities for the dissemination of the RAG's standardization work. Since 2012, when the dictionary was published online, the number of queries has increased continuously. Furthermore, the work of the dictionary is not limited to this, it also offers entertainment and didactic content. In addition, these same resources have been employed to inform about the

standard language, and especially, to establish a closer relationship with speakers, students and professionals. The main objective has been on the one hand, to learn about their needs and adjust the dictionary's responses in relation to them, and on the other, to improve attitudes about the Galician language, for let's not forget that we are dealing here with a minoritized language which is right now in a situation of diglossia. In his article, **Vincent Rivière**, head of linguistic projects at *Lo Congrès Permanent de la Lengua Occitana*, introduces a tool to conjugate Gascon verbs for a broad range of users. This is a significant innovation, as this type of tool did not exist for the Gascon language before. **Gilles Souvay**, member of the research laboratory *Analyse et Traitement Informatique de la Langue Française* (ATILF) engages in his article with the electronic dictionaries developed at the laboratory. After producing the electronic version of the *Trésor de la Langue Française*, the ATILF has developed a number of lexicographic resources: dictionaries (modern, historical, etymological), textual corpus data and other resources and tools for the automatic treatment of language.

Lastly, the fourth section of the book engages with the linguistic and social resources of the digital environment for the standardization of Basque. **Ainhoa Arregi**, lexicographer at UZEI and corresponding member of Euskaltzaindia, presents us in her article with the results of an exercise performed to study the health of the Basque language in the digital environment, taking as its starting point some indicators related to Basque. Not only is normativization important, but also that is quickly and adequately disseminated and, Arregi highlights, the responsibility for this lies with language professionals. **Josu Aztiria**, director of the Language and Technology Department at the Elhuyar Foundation, outlines in his article the windows of opportunity that are currently open for the Basque language in the era of artificial intelligence, and describes the progress made from symbolic strategy to automatic learning. He also proposes a number of fields of intervention in which to get to work. In the following article, **Olatz Arregi**, **Iñaki Alegria** and **Xabier Artola**, lecturers at the UPV (University of the Basque Country) and members of the IXA group, discuss standardization and linguistic technologies, and engage both with the lessons learned so far and the need for planning. Computers, digitalization, Internet, electronic books, smart phones and social media among other things, have all entered our homes and businesses, and, together with these, language technologies have also become part of our everyday lives. Basque began its process of adaptation to the technological context, and technology started to engage with Basque quite early on. And yet, if we compare its particular situation with that of other languages with larger resources behind them, it is still in a state of «incompleteness». And so, the challenges for the future are immense. In their article, the authors propose the direction and main lines of development that the required planning should

take. The last article in this final section is authored by **Jean-Baptiste Battittu Coyos**, member of Euskaltzaindia and in charge of the Tutelary Section of the Academy. In the article, Coyos presents Euskaltzaindia's new resource, called *Basque eskuz esku* (EEE), the aim of which is to provide help for a better use of standard Basque in continental Euskal Herria.

The rationale behind it is both to make the particularities of the area visible, and, alongside it, to introduce Euskaltzaindia's rules in a simplified way, via a website designed specifically with this objective in mind.

Now that we are familiar with all of these experiences, the time has come to reflect, since we now have the tools necessary to go searching for the answer to the question introduced at the beginning of this text. This is going to be the main challenge for the near future.