

# ***Le lexique de l'espéranto : origine, stratégie et évolution***

*Philippe Planchon*

UNIVERSITÉ DE TOURS, DÉPARTEMENT SDL,  
LABORATOIRE LLL (UMR 7270)

philippe.planchon@univ-tours.fr

## **Résumé**

À partir d'une classification générale des langues construites (langues auxiliaires internationales ; langues *a priori* ou *a posteriori* ; schématiques ou naturalistes), l'article envisage la situation particulière de l'espéranto, qui est passé de l'état de projet de langue à celui de langue vivante. La prise en compte du domaine lexical permet de voir les choix opérés par le concepteur initial de la langue, notamment en termes de convergence et de divergence étymologique, et par suite, de comprendre les potentialités d'évolution offertes aux locuteurs. Nous proposons d'interpréter cette incomplétude initiale de façon positive, comme le résultat d'une stratégie pensée afin de rendre possible l'évolution ultérieure de la langue. Les tendances qui s'efforcent depuis lors de réorganiser le lexique (afin d'augmenter la régularité interne de la langue, ou bien de préserver la fidélité aux langues sources) sont le témoignage de cette position d'équilibre instable initialement souhaitée par le concepteur de la langue.

## **1. Introduction**

Comment forger un lexique pour une langue construite ? Cette question, d'ordre très général, mérite d'être posée, notamment dans le cas de l'espéranto. En effet, l'espéranto constitue un cas particulier, dans l'histoire des langues construites, dans la mesure où c'est la seule langue ayant atteint un niveau conséquent d'usage et de diffusion, donnant naissance à une communauté significative de locuteurs sur une longue période de temps (plus de 135 ans). Si l'on considère les langues construites à des fins de communication internationale (que l'on appelle également *langues auxiliaires internationales*), des centaines de projets de langues ont été publiés, et seule

une petite dizaine ont eu un usage réel<sup>1</sup>, mais extrêmement limité (par le nombre de locuteurs, la diffusion et la durée d'usage). Detlev Blanke parle à leur sujet de langues semi-planifiées, réservant alors l'expression de *langue planifiée* au seul cas de l'espéranto<sup>2</sup>. A partir d'une série de 28 étapes permettant de juger de l'extension et du degré de socialisation d'une langue construite<sup>3</sup>, Detlev Blanke observe que l'espéranto est la seule langue qui a atteint les derniers échelons de cette série d'étapes. Il soulève du même coup la question du "succès relatif de l'espéranto" (suivant le titre de l'article de Blanke 2009) par rapport à d'autres projets de langues, qui n'ont pas eu le même destin. Qu'est-ce qui peut expliquer ce succès de l'espéranto ? Est-ce que la planification initiale de la langue a joué un rôle dans ce succès ? Est-ce que le lexique a été un élément déterminant pour cette planification ? Quelle stratégie a présidé à l'élaboration du lexique ?

Face à ces questions, il convient tout d'abord de revenir sur les facteurs possibles qui ont pu jouer dans le succès (relatif) de l'espéranto. Certains auteurs, par exemple Sikosek (2003), considèrent que les circonstances historiques et les conditions sociologiques de réception de l'espéranto sont les principaux facteurs pouvant expliquer que l'espéranto (et non d'autres langues construites) ait reçu un tel accueil et ait pu se développer. Cette vision est cependant trop restrictive, car les facteurs linguistiques ne peuvent être totalement écartés. Ils ne sont certes pas les seuls à avoir joué, mais si Zamenhof avait proposé un projet de langue totalement différent (par son contenu linguistique), par exemple s'il avait proposé un projet de langue similaire au Klingon (créé en 1985 par Marc Okrand), à l'Ithkuil<sup>4</sup>, ou d'une façon plus plausible : une langue philosophique (comme bien des langues proposées avant son époque), tout laisse à penser que l'espéranto n'aurait pas pu prospérer et n'aurait pas été pratiqué jusqu'à aujourd'hui. Cela peut sembler un argument de principe, et c'est pourquoi nous examinerons brièvement le cas des langues philosophiques, afin de montrer en quoi les choix opérés par Zamenhof répondent à une stratégie réfléchie (en particulier au regard de langues construites comme le volapük, dont il avait eu connaissance avant de publier son premier ouvrage sur l'espéranto). Enfin, un troisième ordre de facteurs, au croisement des facteurs externes (sociologiques) et internes (linguistiques), doit être pris en considération. On peut les désigner comme facteurs sociolinguistiques, au sens de la prise en compte des conditions sociologiques d'acceptation d'un tel projet dans

---

<sup>1</sup> Parmi celles-ci, on peut mentionner le Volapük, l'Ido, l'Occidental-Interlingue, l'Interlingua, le Latino Sine Flexione, mais aussi plus récemment : le Lojban, le Kotava, la Lingua Franca Nova – parmi quelques autres langues.

<sup>2</sup> « To understand fully the phenomenon of a truly functioning planned language, we must make a distinction between a project and a language. Scholars often ignore this distinction, attributing to a planned language characteristics that are valid only for a project » (Blanke 2004: 6).

<sup>3</sup> La liste des critères utilisés par Detlev Blanke est citée et commentée en détail par Ida Stria (2015: 78-80), à laquelle nous reprenons cette analyse.

<sup>4</sup> John Quijada, l'inventeur de l'Ithkuil (en 2004), s'est donné pour objectif de concevoir la langue la plus complexe possible, afin de tester l'hypothèse Sapir-Whorf. Il intègre ainsi pour cette langue un système de 96 cas grammaticaux, une morphologie consonantique sur le modèle des langues afro-asiatiques, le système aspectuel des langues bantoues, l'ergativité du basque, le système enclitique des langues kwakiutl, la morphophonologie des langues abkhazes, etc.

l'élaboration même des propriétés linguistiques du projet de langue. Tous ces facteurs jouent un rôle, dont l'exacte mesure est certainement difficile à établir, mais nous nous concentrerons ici sur les propriétés linguistiques de la langue, en tenant compte de l'objectif poursuivi par Zamenhof, dans la mesure où cet objectif a en grande partie guidé ses choix.

## **2. Les différentes stratégies déployées dans les langues construites**

En quoi les choix opérés par l'espéranto dans la constitution de son lexique divergent-ils des choix opérés pour d'autres langues construites ? En suivant une tradition bien établie de l'interlinguistique (Couturat & Leau 1903, Monnerot-Dumaine 1960), nous prendrons tout d'abord en compte la différence entre langues *a priori* et langues *a posteriori*. Le lexique des langues *a priori* est construit indépendamment du lexique de langues déjà existantes. Il ne peut donc être question d'emprunts. On peut considérer ici le cas des langues philosophiques<sup>5</sup>, comme caractéristique d'une telle stratégie. Dans une langue philosophique, tout choix doit être rationnel, et il n'est donc pas question de choisir de façon aléatoire les unités du lexique. Mais cette stratégie exclut de recourir à l'exemple des langues naturelles (perçues comme imparfaites par ces philosophes). Par conséquent, le choix d'une racine doit être entièrement motivé, ce qui conduit à attribuer à chaque unité de forme (lettre ou phonème) une valeur sémantique. Celle-ci est à son tour déterminée par un tableau de la réalité où chaque phénomène est classé, sous la forme d'une taxinomie (Slaughter 1982).

Un mot est donc décomposable en unités minimales de forme et de sens. Comme l'explique Eco, à propos de la langue philosophique de Wilkins (créée en 1668) : « Si (*De*) signifie Élément, alors (*Deb*) doit signifier la première différence ; laquelle (selon les Tableaux) est le Feu ; et (*Deba*) dénotera la première Espèce, qui est la Flamme. (*Det*) sera la cinquième différence sous le Genre, qui est le Météore qui Apparait ; (*Deta*) la première Espèce, c'est-à-dire l'Arc-en-ciel ; (*Deta*) la deuxième, c'est-à-dire le Halo » (Eco 1994: 280).

Les inconvénients de telles langues sont multiples. Les proximités de sens sont systématiquement corrélées à des proximités de formes, ce qui constitue une surcharge considérable en termes d'attention et de mémorisation. Mais cela va plus loin, puisque le choix d'attribuer un sens à chaque phonème conduit à transgresser le principe de double articulation du langage (Martinet 1960). La stratégie des langues philosophiques est de viser une forme de perfection, qui rend en réalité ces langues inutilisables pour des êtres imparfaits comme le sont les êtres humains.

Cet exemple des langues philosophiques montre dès le départ qu'il convient d'accepter des compromis, afin d'obtenir une langue auxiliaire qui soit (certes) plus régulière que les langues nationales, mais qui soit également une langue utilisable en pratique, malgré ses imperfections. Cet équilibre est certainement

---

<sup>5</sup> Pour citer un exemple de langue *a priori* qui n'est pas une langue philosophique, mentionnons le Kotava (créé en 1978 par Staren Fetcey), auquel ne s'appliquent pas plusieurs des critiques que nous formulons par la suite.

un des enjeux de tout projet de langue construite, si celle-ci vise à être effectivement parlée.

De ce point de vue, les langues *a posteriori* procèdent d'une stratégie beaucoup plus raisonnable, puisque les éléments du lexique se trouvent directement empruntés à des langues déjà existantes. Mais là aussi, deux stratégies adverses peuvent être poursuivies. Suivant la première, illustrée notamment par le Volapük (créé en 1879 par Johann Martin Schleyer), l'internationalité serait entravée par l'emprunt direct d'un mot à une langue donnée. Il n'est donc pas nécessairement souhaitable que l'on puisse reconnaître l'étymologie d'un mot<sup>6</sup>. De plus, le Volapük dispose d'une morphologie flexionnelle très riche et synthétique, puisqu'une forme verbale par exemple peut comporter toute une série de préfixes et désinences (marque de diathèse, désinence de temps préfixée, marque de personne, désinence de mode suffixé). Par exemple, *j'aurais été aimé* se traduit par *pilöfoböv* (p – i – löf – ob – öv : passif – plus-que-parfait – radical – 1SG – conditionnel).

Pour composer avec cette complexité morphosyntaxique, le concepteur du Volapük a cherché à limiter la longueur des radicaux, ce qui explique l'existence d'un nombre important de radicaux monosyllabiques. Par ailleurs, pour garantir l'adjonction d'éléments préfixés et suffixés sans hiatus ni cluster, la structure syllabique CVC a été adoptée pour les radicaux, en cherchant là aussi à éviter de déroger à une telle règle. Cela explique par exemple le nom de la langue, formé de *vol* "monde" (< angl.<sup>7</sup> *world*) et de *pük* "langue" (< angl. *speak*). Cela aboutit à déformer un très grand nombre de racines empruntées, afin de respecter ces contraintes morphologiques, ainsi que pour tenir compte de contraintes phonologiques particulières : évitement du /r/ au profit de /l/ (jugé plus facilement prononçable par les locuteurs de langues asiatiques), neutralisation de l'opposition de voisement pour les affriquées et chuintantes, etc. La physionomie très particulière du Volapük a suscité d'innombrables critiques dès son apparition, et même si le fonds lexical a été effectivement emprunté à des langues déjà existantes, il résultait d'un nombre très important de déformations<sup>8</sup>, ce qui a amené à considérer qu'il s'agissait d'un type de langue *schématique*<sup>9</sup>.

À l'extrême opposé de cette stratégie, le concepteur d'une langue peut viser à altérer aussi peu que possible les racines empruntées aux langues sources. Un

---

<sup>6</sup> Comme l'indique par exemple Kerckhoffs (1886: 63) : « Il n'est pas nécessaire de faire remarquer que les radicaux du volapük n'ont rien de commun avec les racines proprement dites des langues auxquelles ils sont empruntés ».

<sup>7</sup> Les abréviations suivantes seront ici utilisées pour les noms de langues mentionnées dans l'article : angl. (anglais), all. (allemand), fr. (français), gr. (grec ancien), ital. (italien), lat. (latin), pol. (polonais), ru. (russe), sl. (langues slaves).

<sup>8</sup> Voici d'autres exemples de racines rendues "méconnaissables" : *nim* (< animal), *fum* (< fourmi), *bel* "montagne" (< all. *Berg*), *tut* "dent" (< angl. *tooth*), *plim* (< compliment)...

<sup>9</sup> Voici un autre exemple de processus rendant méconnaissable la langue source : il s'agit de la formation du lexique du Lojban (créé en 1987 par le Logical Language Group) qui s'appuie sur un calcul statistique (Cowan 1997: 53 ; 71). Un algorithme a été appliqué à six langues sources (anglais, chinois, arabe, russe, hindi, espagnol) pour générer les formes qui reprenaient les phonèmes les plus fréquents pour chaque mot traduit dans ces langues (avec un système de pondération sur les langues, pour tenir compte du nombre de locuteurs de ces langues). Il suffit toutefois de consulter le vocabulaire de base de la langue (1350 mots simples) pour observer que cette stratégie a neutralisé l'origine de ces mots, qui n'est presque jamais reconnaissable (ce qui est là aussi mis en avant comme un avantage de la langue).

exemple caractéristique de cette stratégie est l'Interlingua (créé en 1951 par Alexander Gode). Le souci de fidélité aux langues sources (en l'espèce, aux langues romanes), vise à rendre la langue construite aussi familière et transparente que possible aux locuteurs de ces langues romanes. Ce souci va si loin qu'il intègre des particularismes orthographiques ou de prononciation qui peuvent faire obstacle à la maîtrise d'une telle langue. Par exemple, une même lettre peut avoir différentes prononciations ("c" est prononcé /k/ dans *cavallo*, /ts/ dans *cento*, /ʃ/ dans *chassar...*), et un même phonème peut s'orthographier diversement (/k/ est écrit "c" dans *cavallo*, "k" dans *kilogramma*, "q" dans *quattro...*). Ce souci étymologique va si loin qu'une forme comme *philosophia* est finalement choisie en Interlingua, alors que la graphie *filosofia* se trouve pourtant employée en italien et en espagnol. Ces choix vont donc à l'encontre de l'objectif d'internationalité et de régularité qui devrait convenir pour une langue auxiliaire internationale, mais il donne le sentiment d'une plus grande proximité avec des langues naturelles, ce qui a amené à parler dans de tels cas de langues *naturalistes*.

Face à cette alternative (langues schématiques *vs* langues naturalistes), la question se pose de situer l'espéranto. De fait, on peut légitimement hésiter sur le classement de l'espéranto, dont on souligne tantôt le caractère schématique, tantôt le caractère de langue naturaliste. On parle parfois à ce sujet de langue semi-schématique (Monnerot-Dumaine 1960), ou de langue mixte (Couturat & Leau 1903 ; Janton 1977). Comme nous allons le voir, l'espéranto tente en effet de concilier ces deux exigences, qui peuvent se révéler parfois contradictoires, entre la volonté de conserver l'étymon sans altération dans la langue et la nécessité de l'adapter autant que possible au système de la langue. Nous parlerons dans le cas présent d'une voie médiane. Cette position d'équilibre vise à concilier ces tendances contradictoires, en laissant ainsi la langue se déployer par la suite de façon équilibrée, ou en laissant ouvertes différentes évolutions possibles pour la langue.

### **3. La stratégie mise en œuvre dans l'élaboration de l'espéranto**

Afin de mieux situer cette stratégie, il peut être utile de mettre en perspective ce qui a été proposé par Zamenhof<sup>10</sup> et ce qui constitue aujourd'hui l'usage habituel de l'espéranto<sup>11</sup>. Nous prendrons quelques exemples dans différents niveaux de fonctionnement de la langue, avant de centrer plus spécifiquement notre propos sur le domaine lexical.

L'espéranto est une langue qui distingue clairement le niveau des morphèmes et le niveau des phonèmes (double articulation du langage). Au niveau phonologique, on peut dégager un système de 28 phonèmes par le recours à

---

<sup>10</sup> Nous nous appuyons essentiellement sur le *Fundamento de Esperanto* (1905), qui reprend les fondements de la langue posés initialement par la *Unua Libro* (1887) et par le lexique développé durant cette première période de fonctionnement de la langue (1887-1905). Pour ce qui concerne les étapes préparatoires à la publication de la *Unua Libro*, et notamment les versions antérieures de la langue (*pra-esperantoj*), voir Waringhien (1959) et Mattos (1987).

<sup>11</sup> Nous examinerons celui-ci notamment à partir du corpus en ligne *Tekstaro* (disponible sur [www.tekstaro.com](http://www.tekstaro.com)).

des paires minimales (Wells 1989: 18). La graphie de l'espéranto est régulière, et recoupe en grande partie les conventions de l'API, hormis pour les 7 phonèmes suivants qui s'écrivent différemment : /tʰ/ qui est écrit "c" ; /tʃ/ écrit "ĉ" ; /dʒ/ écrit "ĝ" ; /x/ écrit "ĥ" ; /z/ écrit "ĵ" ; /ʃ/ écrit "ŝ" ; /w/ écrit "ŭ". L'avantage de ce système graphique est d'éviter les digrammes, tout en conciliant des pratiques divergentes. Par exemple, l'usage d'accents sur les consonnes pour transcrire des chuintantes est caractéristique de systèmes orthographiques en lituanien et dans des langues slaves écrites en alphabet latin (polonais, tchèque, slovaque, croate...), tandis que la forme de l'accent s'apparente au circonflexe du français<sup>12</sup>.

Un certain sens pratique conduit pourtant Zamenhof à proposer une option de transcription, introduite dès 1888 (*Dua Libro*), pour remplacer l'accent circonflexe par un digramme lorsque des contraintes matérielles (typographies, imprimerie) ne permettent pas l'usage du circonflexe sur des consonnes. De sorte que ĉ peut être écrit *ch*, ĝ peut être écrit *gh*, et ainsi de suite.

Comme le montrent ces exemples, l'espéranto est initialement conçu comme le résultat de compromis, afin de concilier des exigences potentiellement contradictoires.

C'est également en ce sens que l'on peut comprendre l'introduction d'un suffixe *-um* avec un sens indéterminé, alors que tous les autres affixes de la langue possèdent un sens clairement défini<sup>13</sup>. Un tel suffixe a pu paraître comme une anomalie à beaucoup de locuteurs au début de l'histoire de la langue, mais il vient en réalité remplir des fonctions spécifiques (Cherpillod 2011), et s'est même révélé productif dans l'usage de la langue (Gončarov 2021). Un phénomène similaire peut être observé sur le plan syntaxique avec la préposition *je*, qui là aussi fonctionne sans posséder de sens précis, mais comme l'indique le *Fundamento* : « *je* se traduit par différentes prépositions ; sa signification est toujours aisément suggérée par le sens de la phrase »<sup>14</sup>. À la différence du suffixe *-um*, cette préposition est beaucoup moins productive, et son emploi s'est spécialisé avec certaines expressions (par exemple : *Je via sano !* "À votre santé !"), certaines fonctions (par exemple pour indiquer l'heure : *mi venos je la dek-kvina* "je viendrai à 15 heures") ou avec certains verbes (par exemple : *fidi je iu* "avoir confiance en quelqu'un").

Ce qui aura pu passer pour une forme d'imprécision ou d'indétermination devrait être plutôt interprété comme une stratégie visant à faciliter l'entrée dans l'usage de la langue. Ces procédés sont mis à disposition des locuteurs pour faciliter l'évolution de la langue. Il peut donc s'agir de simples possibilités,

---

<sup>12</sup> Waringhien (1980: 90) montre par des exemples comment ce choix concilie les exigences d'une écriture phonétique, d'un système graphique apparenté aux langues slaves, et d'une préservation des étymons (issus des langues romanes) ; par exemple : *ĵurnalo* /ʒur'nalo/ (journal), *ĝardeno* /dʒar'deno/ (jardin), etc.

<sup>13</sup> « Si nous devons utiliser un suffixe, mais si le sens ne nous montre pas quel suffixe nous devons prendre, nous utilisons le suffixe *-um* », *Dua Libro* (1888) - ma traduction.

<sup>14</sup> Les références au *Fundamento de Esperanto* (Zamenhof 1905) s'appuient sur l'édition critique de l'ouvrage, publiée en 1991. Dans le cas présent, l'explication sur cette préposition est donnée dans le vocabulaire du *Fundamento* (s.v. *je*).

offertes aux utilisateurs de la langue, mais qui pourront prospérer ou rester lettre morte.

Par exemple, la possibilité de doubles consonnes est envisagée dès le départ, puisque l'on a la forme *Ŝillero* (et non *Ŝilero*) comme espérantisation du nom propre *Schiller* dans la *Unua Libro*. De même, Zamenhof n'hésite pas à utiliser dans les premiers textes de la langue des formes comme *getto* (ghetto) ou *villo* (villa). Il n'est donc nullement exclu de généraliser en espéranto l'emploi de doubles consonnes (qu'il resterait toutefois à prononcer différemment des consonnes simples, afin de respecter la caractère phonologique de l'écriture). Toutefois, il ne s'agit que d'une potentialité, et dans le cas présent, celle-ci n'a pas été mise à profit dans le développement de la langue. Certaines formes peuvent présenter ces doubles consonnes (et cela encore aujourd'hui, par exemple dans le nom de pays *Finnlando*, construit sur le substantif *finno* "Finnois"), mais celles-ci sont évitées autant que possible, et il ne s'agit donc pas d'un phénomène productif.

De même, l'usage de l'élision à la place de la désinence du substantif (-o) a été proposé dès l'*Unua Libro* (comme règle de base de la langue), et a pu être employé dans un premier temps par de nombreux locuteurs (y compris dans des textes courants, ou dans la prose). Toutefois, cette possibilité a été peu à peu délaissée, et n'existe plus aujourd'hui que dans des textes poétiques ou des chansons. Cette potentialité n'a donc pas été productive dans l'évolution ultérieure de la langue.

À l'inverse, un phénomène comme celui de l'usage autonome des morphèmes s'est révélé très productif dans la langue<sup>15</sup>, même s'il n'a été timidement introduit que peu à peu<sup>16</sup>. L'usage autonome des morphèmes est un principe potentiellement contenu dans la langue, puisque celle-ci est conçue de façon agglutinante (et présente également des aspects isolants). Tout élément morphologique possède une forme d'autonomie de principe (autonomie de sens et de forme) au regard des éléments auxquels il s'adjoint. Par suite, un suffixe possède un sens propre et une forme stable (sans allomorphe) quel que soit le radical auquel il est ajouté. Cette stabilité lui permet donc de fonctionner également comme radical pour former de nouveaux mots, ce qu'on peut définir comme l'usage autonome des morphèmes. Par exemple, à partir du suffixe *-ebl* (*lav-ebl-a* "lavable" ; *fleks-ebl-a* "flexible" ; *kompren-ebl-a* "compréhensible"), il est possible de former un adverbe par adjonction de la désinence *-e* au suffixe *-ebl* utilisé comme radical (*ebl-e* "possiblement, peut-être"). Et ce nouveau radical peut à son tour être utilisé comme base de dérivation (*ebligi* "rendre possible", *ebleco* "possibilité", *malebla* "impossible").

Il ne s'agit plus alors d'un suffixe au sens strict, et l'on a affaire à un phénomène de transcatégorialité généralisée. En espéranto, celle-ci est susceptible

---

<sup>15</sup> Non seulement de nouvelles formes continuent d'entrer en usage, mais les usages autonomes des morphèmes sont souvent préférés par les locuteurs (par rapport à d'autres formations équivalentes). Par exemple, la forme *eta* (basée sur le suffixe diminutif *-et*) peut facilement venir remplacer la forme *malgranda* "petit" (obtenue par antonymie de *granda* "grand"). Le corpus *Tekstaro* ([www.tekstaro.com](http://www.tekstaro.com)) montre un emploi très limité de *eta* par rapport à *malgranda* pour la période 1887-1954 (8,2%), mais un emploi en très nette progression pour la période 1954-2024 (47,5%).

<sup>16</sup> À titre d'illustration, voici les dates des premiers usages autonomes de plusieurs suffixes : *ebl* (1888), *emo* (1889), *eksa* (1889), *estro* (1890), *igi* (1891), *eta* (1892), *ino* (1893), *ilo* (1894), *iĝi* (1896)...

d'opérer quel que soit l'élément pris en considération (suffixe, préfixe, préposition, corrélatif, particules, etc.)<sup>17</sup>. Ce fonctionnement spécifique de l'espéranto avait été explicité dès le départ<sup>18</sup>, mais les implications d'un tel principe morphologique n'ont été véritablement perçues qu'au fur et à mesure de la généralisation de ces formes autonomes, lorsque celles-ci sont entrées réellement dans l'usage des locuteurs.

Si l'on peut résumer ce que nous venons d'envisager à travers ces quelques exemples (en phonologie, syntaxe et morphologie), la stratégie adoptée par Zamenhof n'était pas de fixer la langue de façon rigide dès le commencement, mais bien de rendre possible ou de faciliter une évolution naturelle de la langue, en y inscrivant différentes potentialités.

#### 4. Stratégie d'élaboration de la langue sur le plan lexical

Nous en venons à présent à la composante lexicale de la langue. Celle-ci est envisagée par Zamenhof suivant la même stratégie que nous avons exposée jusqu'à présent. Il s'agit pour l'initiateur de la langue de rendre possible l'évolution future de celle-ci, tout en assurant une stabilité suffisante à la langue dans sa forme initiale.

Par suite, il développe l'idée d'une évolution continue du lexique par une méthode qu'il expose comme la méthode des néologismes et des archaïsmes : « nous avons seulement besoin de créer un certain nombre de nouveaux mots, qui (...) ne remplaceront pas les anciens mots inopportuns, mais ils existeront seulement parallèlement à ceux-ci, en tant que synonymes, lesquels existent certes en grand nombre en toute langue – en tant que néologismes, qui vraisemblablement expulseront peu à peu les anciennes formes et les rendront archaïques, cependant jamais ils ne les annuleront ni n'enlèveront leur valeur » (*Circulaire de Zamenhof adressée au Comité Linguistique* (1906), citée par Pabst (2015: 69) ; ma traduction). Ou encore, dans l'avant-propos du *Fundamento* : « Avec le temps la nouvelle forme peu à peu expulsera l'ancienne forme, qui deviendra un archaïsme, comme nous voyons ceci se produire dans toute langue naturelle » (ma traduction).

Comme l'explique Detlev Blanke, Zamenhof avait en vue l'articulation complexe qui peut exister entre le code linguistique et l'usage qui en est fait : « Zamenhof intuitively understood the essential problems of structure and functioning of a language, among them the relation between linguistic system (*langue*) and its concrete application (*parole*), the importance of the language community, the relationship between stability and evolution, and the problem of norms and their codification. » (Blanke 2009: 255).

Dans le cas présent, cela signifie qu'il n'y a pas de fétichisme du mot chez Zamenhof (par contraste par exemple avec la position de Schleyer à l'égard du

---

<sup>17</sup> Par exemple : préfixe *eks-* (*eks-direktoro* "ex-directeur") / adjectif *eksa* ("ayant quitté ses fonctions") ; préposition *per* ("au moyen de") / verbe *peri* ("servir de moyen") ; corrélatif *ĉiam* (*ĉi-am* : collectif-temps "toujours") / adjectif *ĉiama* ("éternel") ; particule *ne* (marqueur de négation) / verbe *nei* ("nier").

<sup>18</sup> « Ce qu'on appelle désinence ou préfixe ou suffixe, est un mot entièrement autonome [en espéranto], qui toujours conserve une signification identique, et peu importe s'il est utilisé à la fin ou au début d'un autre mot ou de façon indépendante, de sorte que chaque mot peut être tout aussi légitimement utilisé comme mot radical ou comme élément grammatical. » (Zamenhof, *Unua Libro*, 1887: 6).

volapük). Ce n'est pas le mot singulier qui lui importe, mais l'inscription de ce mot dans un système linguistique qui doit conserver une forme de stabilité pour pouvoir se développer comme langue vivante.

De fait, un bon nombre de mots introduits dans la langue durant cette première période (1887-1905) sont tombés depuis lors dans l'archaïsme, ainsi que le prévoyait Zamenhof, ne serait-ce qu'en raison de l'évolution technique, scientifique et sociale<sup>19</sup>.

Inversement, de nombreux néologismes sont apparus du fait de l'évolution de nos sociétés, néologismes qui bien entendu n'existaient pas à l'époque de Zamenhof<sup>20</sup>.

En d'autres termes, le lexique conçu à l'origine n'est pas vu par l'auteur lui-même comme ce qui devrait constituer à terme le lexique total de la langue, bien au contraire ! Il s'agit en réalité d'un ensemble d'illustrations de principes de formation des mots en espéranto.

L'une des "règles de grammaire"<sup>21</sup> de la langue est d'ailleurs consacrée à la question des emprunts, à savoir la règle 15 : « Les mots "étrangers" c.-à-d. ceux que la plupart des langues ont empruntés à la même source, ne changent pas en Esperanto. Ils prennent seulement l'orthographe et les terminaisons grammaticales de la langue. Mais quand, dans une catégorie, plusieurs mots différents dérivent de la même racine, il vaut mieux n'employer que le mot fondamental, sans altération, et former les autres d'après les règles de la langue internationale. Ex.: tragédie – tragedi'o, tragique – tragedi'a. » (*Fundamento*). On peut relever dans cette règle plusieurs objectifs, qui sont en réalité constants chez Zamenhof, à savoir : l'internationalité (la "plupart des langues" / la "même source"), la possibilité de faire évoluer la langue (par des emprunts), et le souci de cohérence du système linguistique (puisque'il convient de respecter les règles de la langue).

## **5. Utilisation de formes internationales et convergentes**

Considérons tout d'abord la notion d'internationalité. Zamenhof considère à juste titre un phénomène qui n'a cessé de s'amplifier depuis son époque, à savoir la diffusion d'un lexique international emprunté par les langues les plus diverses. Ces mots "étrangers" entrent en réalité de plein droit dans le lexique de l'espéranto, dans la mesure où l'espéranto vise à servir de langue internationale. Ce lexique international est composé par exemple de mots savants, d'origine grecque ou latine (*biologio* "biologie", *oksigeno* "oxygène", *energio* "énergie", *patelo* "rotule", *diagnozo* "diagnostic"...). Toutefois, ces

---

<sup>19</sup> Voici quelques exemples de mots qui sont (tout autant d'ailleurs en espéranto qu'en français) devenus des archaïsmes, et ne sont généralement pas connus des locuteurs de l'espéranto (ou du français) aujourd'hui : *gado* ("merluce"), *kitelo* ("souquenille"), *nankeno* ("nankin"), *stupo* ("étoupe"), *tombako* ("tombac"), *surtuto* ("redingote")...

<sup>20</sup> Parmi les exemples les plus évidents, voici quelques mots d'usage actuel et apparus bien après la mort de Zamenhof : *televido* ("télévision"), *raketo* ("fusée"), *fridujo* ("réfrigérateur"), *komputilo* ("ordinateur"), *interreto* ("internet"), etc.

<sup>21</sup> La dimension "grammaticale" de cette règle concerne uniquement l'application des flexions (et donc la morphosyntaxe), même si, par la force des choses, cette règle sert également à montrer à l'utilisateur de la langue la façon dont la langue peut s'enrichir de nouveaux mots (ce qui concerne du même coup la dimension lexicale de la langue).

mots ne sont pas nécessairement apparus durant l'antiquité, mais reflètent aussi les pratiques terminologiques des sciences modernes. Par ailleurs, le vocabulaire de la botanique ou de la zoologie est marqué par l'usage récurrent de racines latines, au point que les classifications et nomenclatures traditionnelles prennent la forme du latin, ce qui explique du même coup le privilège accordé à l'emprunt de racines latines dans la formation des mots en espéranto, par exemple : *azeno* "âne" (lat. *asinus*), *sciuro* "écureuil" (lat. *sciurus*), *erinaco* "hérisson" (lat. *herinaceus*), *kverko* "chêne" (lat. *quercus*), *hedero* "lierre" (lat. *hedera*), *betulo* "bouleau" (lat. *betulla*), etc. Mais là aussi, cela ne signifie pas que ce vocabulaire international ait pris naissance dans la langue de Cicéron, mais que la forme latinisée de mots étrangers sert de standard. De nombreux noms de plantes, de fruits, d'animaux, ont été internationalisés à partir de noms empruntés aux langues sources les plus diverses. Voici quelques exemples : *tomato* (< *xitomatl* du nahuatl), *ananaso* (< *naná-naná* du tupi-guarani), *kanguruo* (*gangurru* de la langue aborigène guugu yimidhirr), *kaŭĉuko* (< *caochu* du quechua), *tulipo* (< *tülbent* du turc), *maizo* (< *mahiz* de l'arawak)...

Bien entendu, le vocabulaire fondamental de la langue reste à établir, mais ces mots internationaux constituent un fonds important du lexique. Dans d'autres cas, Zamenhof privilégie certaines langues pour le rôle qu'elles jouent ou ont pu jouer comme langues internationales. C'est pourquoi il forme un grand nombre de racines en les empruntant au latin, au français, à l'anglais, et de façon beaucoup plus marginale, à l'allemand, à l'italien, au russe et au grec. Souvent interprété comme le choix de langues européennes, il s'agit en réalité du choix de langues en raison de leur usage comme langues internationales (à l'époque de Zamenhof). Parmi ces langues, le latin reçoit un rôle privilégié par sa neutralité, dans la mesure où il s'agit d'une langue qui n'a plus de locuteur natif. C'est par exemple du latin que l'espéranto reçoit la plupart de ses prépositions (*sub*, *super*, *per*, *pro*, *inter*, *trans*, *de*, *post*...). Par ailleurs, conscient de privilégier alors certaines langues au détriment d'autres langues, Zamenhof déploie des stratégies pour atténuer une trop grande dépendance entre l'espéranto et l'une ou l'autre de ces langues sources<sup>22</sup>.

Pour éviter de marquer l'espéranto par une langue source unique, Zamenhof est ainsi amené à diversifier ses emprunts entre différentes langues (ou familles de langues) dont il a connaissance. C'est particulièrement remarquable pour des exemples de mots fonctionnant comme des sous-systèmes de la langue, par exemple :

- repérage temporel : *hierau* (< fr. *hier*) / *hodiaŭ* "aujourd'hui" (< lat. *hodie*) / *morgaŭ* "demain" (< all. *Morgen*)
- particules : *jes* "oui" (< angl. *yes*), *nu* "eh bien" (< all. *nu*), *ĉu* "est-ce que ?" (< pol. *czy*), *kaj* "et" (< gr. *καί*), *do* (< fr. *donc*), *sed* "mais" (< lat. *sed*)...

---

<sup>22</sup> « L'homogénéité (une seule source) pour tous les mots de la langue n'est pas seulement tout à fait superflue (car pour les utilisateurs de la langue il est tout à fait indifférent de dire "blinda" ou "ceka", "hundo" ou "kano", etc.), mais ce serait même clairement dangereux, car cela serait comme mettre en esclavage tous les peuples par rapport aux peuples romans, ce qui provoquerait à plus ou moins brève échéance une trop grande et juste protestation contre notre langue » (Circulaire de Zamenhof adressée au Comité Linguistique (1906), citée par Pabst (2015: 69) - ma traduction).

- pronoms personnels : *mi* (lat. *mihi*), *ili* (< fr. *ils*, lat. *illi*), *ŝi* (< angl. *she*), *oni* (< fr. *on* / angl. *one*), *si(n)* (< fr. *se* / sl. *si-* / all. *sich*)<sup>23</sup>.

Comme on le remarque ici, certaines formes sont analogues entre les langues. C'est le cas également des pronoms *ni* et *vi* pour lesquels le latin et les langues romanes comportent ces consonnes initiales de façon régulière (le *-i* étant un élément commun aux différents pronoms personnels de l'espéranto).

L'idée de retenir des formes convergentes, qui neutralisent du même coup la provenance du mot, est un principe mobilisé par Zamenhof lorsque cela est possible. Ce principe est naturel dans le cas de mots internationaux, mais peut également se rencontrer entre quelques langues utilisées comme langues sources. C'est par exemple le cas de l'adverbe temporel *nun* "maintenant", qui se trouve résulter d'une étymologie convergente (puisque c'est une forme qu'on peut retrouver en grec (*νῦν*), en latin (*nunc*) et en allemand).

Cette recherche de convergence peut aussi conduire à combiner des étymons pour obtenir une forme que l'on pourra qualifier de forme hybride. Cette hybridation se rencontre dans les langues naturelles avec les mots-valises. Elle présente ici l'intérêt de neutraliser l'origine du mot, tout en rappelant des formes déjà existantes dans les langues naturelles. C'est par exemple le cas d'un certain nombre de formes à initiale *s-*, qui reçoivent l'initiale *ŝ-* /*ʃ*/ par influence de l'allemand (et très probablement du yiddish). Voici quelques exemples (nous reprenons plusieurs de ces exemples de Waringhien 1959: 80)

:

- all. *Stein* [ʃtajn] / angl. *stone* [stəʊn] : *ŝtono* "pierre"
- all. *Sturm* [ʃturm] / angl. *storm* [stɔ:m] : *ŝtormo* "tempête"
- all. *Stufe* [ʃtu:f<sup>9</sup>] / ru. ступень [stupjeɲ] : *ŝtupo* "marche"
- all. *spritzen* [ʃpritsən] / ital. *spruzzare* [sprutsare] : *ŝpruci* "jaillir"

Par ailleurs, l'hybridation peut associer les deux étymons par plusieurs phonèmes partagés, et par plus d'un phonème repris de l'un et l'autre des deux étymons. Il ne s'agit donc plus simplement d'une influence, mais bien d'une combinatoire. Voici quelques exemples représentatifs de ce phénomène pour différentes combinaisons de langues :

- all. *Lawine* / ital. *valanga* : *lavango* "avalanche"
- lat. *speculum* / all. *Spiegel* : *ŝpegulo* "miroir"
- all. *Kuchen* / angl. *cake* : *kuko* "gâteau"
- angl. *chapter* / fr. *chapitre* : *ĉapitro*
- all. *gleiten* / fr. *glisser* : *gliti*
- ru. *stanok* / angl. *stand* / fr. *établi* : *stablo*

## **6. Les cas de divergence et d'altération des formes empruntées**

Ces cas d'hybridation montrent comment utiliser les étymons de plusieurs langues sources afin d'obtenir une forme convergente en espéranto. Un processus inverse consiste à tirer profit de la divergence des langues sources

---

<sup>23</sup> Sur l'étymologie des pronoms personnels et des corrélatifs en espéranto, voir notamment Waringhien (1959: 59 *squ*).

ou de la modification des étymons pour mieux satisfaire des exigences internes à la langue cible.

Ainsi, le recours à des sources étymologiques différentes peut se justifier pour éviter des cas d'homonymies. Par exemple, emprunter les mots "seul", "soleil" et "sol" directement au latin conduirait à des formes synonymes (avec une racine identique *sol-* dans chaque cas de figure). C'est pourquoi "seul" *sola* est emprunté au latin, tandis que "soleil" *suno* est emprunté à l'anglais *sun*, et "sol" *grundo* est emprunté à l'allemand *Grund*. De même, le français *raison* ne peut fournir que l'étymon de *rezoni* "raisonner", mais il faut un autre étymon pour *raison* (au sens de faculté), à savoir le latin *ratio* (donnant *racio* en espéranto), et un autre encore pour *avoir raison* (opposé à *tort*), fourni cette fois-ci par les langues slaves avec la forme *pravi* "avoir raison".

Un autre exemple est fourni par la préposition *post* "après" empruntée au latin, qui bloque du même coup la possibilité d'emprunter *poste* au français, que ce soit pour un poste (fonction occupée) pour lequel la forme *posteno* est empruntée à l'allemand, ou pour la poste (système de courrier) où c'est la forme *poŝto* qui est retenue, après avoir été empruntée aux langues slaves.

Cet enjeu d'évitement des situations d'homonymie conduit Zamenhof à composer avec la forme même des mots, et en particulier à modifier le voisement des consonnes. Cette modification peut ainsi opérer à partir d'un même étymon. Par exemple, *porti* "porter" provient du latin *portare*, mais cela exclut donc d'utiliser pour "porte" le latin *porta* (fr. *porte*, it. *porto*...), d'où le voisement de /t/ en /d/ pour former *porĉo* "porte". Quant au port (it. *porta*), il requiert un autre étymon, emprunté aux langues germaniques, qui donnera *haveno* "port" en espéranto<sup>24</sup>.

Cette altération consonantique se retrouve mobilisée pour différents couples de mots, par exemple : *pesi* / *pezi* (resp. "peser" tr. / "peser" intr.), *akordi* / *agordi* (resp. "accorder" au sens large / "accorder" en parlant d'un instrument de musique), *lavo* / *lafo* (resp. le "lavage" / la "lave"), *versi* / *verŝi* (resp. "versifier" / "verser").

Parmi les altérations que l'on peut relever, il faut également mentionner le choix d'emprunter certains mots (selon les cas) sous leur forme orale ou sous leur forme écrite. La forme orale est la plus fréquente, mais Zamenhof n'exclut pas l'emprunt à la forme écrite du mot (phénomène que l'on observe également dans des langues naturelles), comme par exemple : *vulpo* /vulpo/ "renard" < lat. *vulpes* /wulpes/ ; *soifi* /soifi/ < fr. "avoir soif" /swaf/ ; *birdo* /birdo/ "oiseau" < angl. *bird* /bɜ:rd/. Cette pratique a ainsi pu perdurer, comme en témoignent des mots plus récents, comme *boato* /boato/ "bateau" (< angl. *boat* /bəʊt/), ou *teamo* /teamo/ "équipe" (< angl. *team* /ti:m/).

Une autre possibilité offerte à l'espéranto dans l'adaptation des emprunts est l'évitement des isoméries. Conscient du risque de mécoupure, c'est-à-dire d'une réinterprétation erronée d'un mot comme mot dérivé (alors qu'il s'agit

---

<sup>24</sup> Le choix de deux langues sources différentes ne prémunit pas nécessairement contre le risque d'homonymie, par exemple : en empruntant au latin pour le "lieu" (*loko* < lat. *locum*), l'étymon *locken* "attirer" ne supprime pas le risque d'homonymie, lequel doit alors être contourné par le voisement de la consonne /k/ en /g/ : *logi* "attirer" < all. *locken*.

d'un mot simple), Zamenhof a ainsi proposé de modifier certaines formes<sup>25</sup>. Par exemple, l'emprunt du mot *cigaretto* sur la base du fr. *cigarette* (ou de formes analogues dans d'autres langues), pourrait amener l'interprétation *cigar-et-o* "petit cigare" par la présence apparente d'un suffixe *-et* "diminutif". Pour éviter cette mécoupure, Zamenhof a modifié le radical emprunté au profit de la forme *cigaredo* "cigarette". Voici quelques exemples de ce mécanisme, où la forme est modifiée pour éviter l'interprétation de l'initiale du mot comme présence d'un préfixe (tel que *de-* "provenance" ou *re-* "itération"), ou de la finale du mot comme présence d'un suffixe (tel que *-et* "diminutif", *-in* "féminin" ou *-il* "outil") :

À l'initiale du mot :

- *di* au lieu de *de* : *diboçi* "agir en débauché", *difini* "définir"
- *ri* au lieu de *re* : *ripeti* "répéter", *rikolti* "récolter"

À la finale du mot :

- *ed* au lieu de *et* : *cigaredo* "cigarette", *planedo* "planète"
- *en* au lieu de *in* : *karabeno* "carabine", *mateno* "matin"
- *el* au lieu de *il* : *mirtelo* "myrtille", *makzelo* "mâchoire" (< lat. *maxilla*)

Une autre façon d'éviter l'isométrie est de supprimer la partie du mot posant difficulté. Un tel procédé de troncation est également illustré par divers exemples. Ainsi, pour éviter *-in* "féminin", la forme retenue est *intesto* "intestin" (< lat. *intestinum*), de même que *prokrasti* "procrastiner" est la forme tronquée à partir du latin *procrastinare*. Pour éviter *-it* (marque du participé passé), il y a également apocope du lat. *sagitta* pour former *sago* "flèche". Et cela peut également affecter le début du mot, comme l'aphérèse de "déchirer" pour former *ŝiri* (afin d'éviter la mécoupure d'un préfixe *de-*). On peut d'ailleurs noter que de telles troncations sont aussi mobilisées pour des mots jugés trop long lorsqu'ils sont empruntés, par exemple *nepre* "surtout" (< ru. *neprejmenno*), ou encore *hidrargo* "mercure" (< lat. *hydrargyrus*).

Ces différentes déformations contribuent certainement à neutraliser l'origine des mots, et donc à intégrer plus facilement ces racines dans le fonds propre de la langue. Toutefois, ces déformations n'interviennent pas au hasard, et encore moins par principe (comme dans les langues schématiques), mais uniquement en réponse à certaines contraintes, de façon mesurée et afin d'assurer une certaine homogénéité dans le lexique de base. Cet équilibre n'est pas parfait. Toutefois, ces différentes solutions ont le mérite de proposer aux utilisateurs de la langue de nouvelles possibilités d'évolution et d'adaptation pour les nouveaux mots susceptibles d'être empruntés et intégrés au cours de l'évolution de l'espéranto.

## **7. L'évolution de l'espéranto depuis l'époque du *Fundamento***

Nous avons jusqu'à présent considéré la première période durant laquelle la langue espéranto a pris forme, à savoir la période allant de 1887 (*Unua Libro*)

---

<sup>25</sup> Sur le risque de mécoupure (ou de métanalyse), nous reprenons ici plusieurs exemples identifiés par Waringhien (1980: 47), même si nous ne partageons pas la même interprétation que lui sur la stratégie lexicale suivie par Zamenhof.

à 1905 (*Fundamento*). Après cette période, il devient possible de raisonner en terme de diachronie<sup>26</sup>. En effet, la langue n'est plus en phase d'élaboration, et toute évolution ultérieure opère de façon progressive, par l'influence de ses locuteurs (qu'il s'agisse de simples locuteurs ou de locuteurs ayant une forme d'autorité sur la langue, tels que les membres de l'Académie d'Espéranto ou les écrivains). La langue entre dans sa vie sémiologique, ainsi que le prévoyait Saussure dans son *Cours de linguistique générale* (cours qui avait justement eu lieu au début du 20<sup>ème</sup> siècle, à l'époque du *Fundamento*) :

« Cette évolution est fatale ; il n'y a pas d'exemple d'une langue qui y résiste. Au bout d'un certain temps, on peut toujours constater des déplacements sensibles. Cela est si vrai que ce principe doit se vérifier même à propos des langues artificielles. Celui qui en crée une la tient en main tant qu'elle n'est pas en circulation ; mais dès l'instant qu'elle remplit sa mission et devient la chose de tout le monde, le contrôle échappe. L'espéranto est un essai de ce genre ; s'il réussit, échappera-t-il à la loi fatale ? Passé le premier moment, la langue entrera très probablement dans sa vie sémiologique (...) L'homme qui prétendrait composer une langue immuable, que la postérité devrait accepter comme telle, ressemblerait à la poule qui a couvé un œuf de canard : la langue créée par lui serait emportée bon gré mal gré par le courant qui entraîne toutes les langues » (Saussure 1995: 111).

Cette évolution se voit tout d'abord dans le nombre de nouveaux mots qui apparaissent entre 1905 et aujourd'hui. Si on considère cette évolution sur le plan quantitatif, on peut ici mentionner quelques ouvrages lexicographiques de référence<sup>27</sup> :

- UL (*Universala Vortaro / Unua Libro*, 1887) : 920 radicaux
- FdE (*Fundamento de Esperanto*, 1905) : 2640 radicaux
- AVE (*Akademia Vortaro / + Oficialaj Aldonoj*) : 4822 radicaux
- PV (*Plena Vortaro*, Grosjean-Maupin 1930) : 6900 radicaux
- PIV (*Plena Ilustrita Vortaro*, Waringhien 1970) : 15200 radicaux
- NPIV (*Nova Plena Ilustrita Vortaro*, Duc Goninaz 2020) : 16780 radicaux<sup>28</sup>

Sur le plan qualitatif, on peut tout d'abord noter qu'une des parties les moins stables de la langue concerne les noms propres, dans la mesure où il s'agit de mots véritablement étrangers à la langue, susceptibles d'être ou non intégrés dans le système phonologique et morphologique de l'espéranto<sup>29</sup>. C'est particulièrement le cas avec les noms de pays, qui subissent une évolution et une adaptation continue depuis cette première période. Par exemple, le radical

---

<sup>26</sup> Sur la répartition de l'évolution de l'espéranto en différentes périodes, voir Koutny (2018), Planchon (2021: 41-42).

<sup>27</sup> Les ouvrages successifs incluent normalement les états lexicographiques antérieurs. Donc le nombre de radicaux indiqué est chaque fois un total qui doit s'entendre comme incluant les radicaux précédemment décomptés.

<sup>28</sup> Les radicaux correspondent aux entrées du dictionnaire. Par ailleurs, le PIV inclut dans le détail des articles 39400 mots, et le NPIV en inclut 46890. Notons que les mots indiqués par les dictionnaires ne sont qu'une petite partie des mots susceptibles d'être formés dans la langue (du fait de ses propriétés agglutinantes, et de son jeu important d'affixes).

<sup>29</sup> Par exemple, pour son étude de la structure syllabique de l'espéranto, Van Oostendorp (1999) exclut ces mots, car il juge que ceux-ci n'appartiennent pas au noyau lexical de la langue.

*ĥin-* (Chine) a été abandonné au profit de *ĉin-*<sup>30</sup> ; le suffixe *-ujo* est progressivement remplacé par *-io* (*Ĉinio* au lieu de *Ĉinujo*). De plus, pour certains noms de pays, c'est l'endonyme qui est aujourd'hui privilégié<sup>31</sup> au lieu de l'exonyme (même si celui-ci possède une plus grande internationalité), par exemple : *Bharato* au lieu de *Hindio* (*Hindujo*) ; puis avec adaptation du radical à la phonologie de l'espéranto : *Barato* au lieu de *Bharato*<sup>32</sup>.

On observe par ailleurs lors de l'apparition de néologismes une phase de concurrence entre différentes formes, suivie d'une stabilisation au profit d'une forme qui se trouve finalement privilégiée par les locuteurs. Les institutions linguistiques (comme l'Académie de l'Espéranto) n'interviennent qu'*a posteriori* pour entériner ces usages, souvent après une période assez longue. L'évolution opère donc de la même façon que pour des langues vernaculaires. On peut ainsi observer la concurrence entre différentes formes lorsque la technologie apparaît ("télévision" : *televizjo*, *televizjono*, *televizoro*, *televido*, *televideo*, *televidilo*...), puis une des formes s'impose (dans le cas présent : *televido*) et sert de base à la formation ultérieure de dérivés (ainsi *televidilo*, formé par adjonction du suffixe *-il* "outil", renvoie à l'appareil de télévision)<sup>33</sup>. Les emprunts ne sont pas de simples calques, mais supposent une adaptation au système de la langue, par exemple *tvitero* ou *vikipedio* sont les formes aujourd'hui usuelles (pour *twitter* et *wikipédia*) où l'on remarque l'emploi de la consonne /v/ (et non /w/) en raison de la phonologie de l'espéranto. De même, le mot *interreto* (composé de *inter-* et *reto*) a été préféré au mot *interneto* (copié de l'anglais), dans la mesure où *reto* ("filet") appartenait déjà à la langue – et s'utilise depuis comme forme abrégée pour *interreto* et comme préfixe (par exemple : *retpaĝo* "page Internet", *retmesaĝo* "message Internet").

Ces évolutions peuvent également affecter des mots déjà entrés dans la langue depuis assez longtemps, par une adaptation plus lente aux structures de l'espéranto. C'est en particulier le cas de ce qu'on appelle les "pseudo-suffixes", qui se trouvent introduits dans la langue sans fonctionner comme suffixes, mais comme traces d'opérations de suffixation qui ont opéré dans la langue source. C'est ainsi le cas pour des mots empruntés aux langues romanes (ou directement au latin). Par exemple, *inaŭguracio* est la forme empruntée du latin *inauguratio* comportant un suffixe *-tio* (action accomplie) qui est inconnu de l'espéranto. Par suite, la forme réduite *inaŭguro* est aujourd'hui plus massivement employée<sup>34</sup>, forme à laquelle on peut d'ailleurs adjoindre de façon régulière un suffixe *-ad* (*inaŭgurado*) si l'on souhaite insister sur l'aspect

---

<sup>30</sup> La forme *ĥinujo* est la plus fréquente dans Tekstaro pour la période 1887-1940, mais elle n'apparaît plus dans la période 1940-2024.

<sup>31</sup> Cette évolution est plus récente, et s'observe notamment dans les usages actuels de l'espéranto. L'étude de Johansson (2015) mentionne la répartition suivante des formes dans son corpus Internet (élaboré en 2004) : *Hindujo* (428), *Hindio* (966), *Bharato* (114), *Barato* (1250).

<sup>32</sup> En effet, ni la séquence /bh/ ni l'aspirée /b<sup>h</sup>/ n'existent par ailleurs en espéranto, d'où l'émergence de la forme /barato/.

<sup>33</sup> Sur l'émergence de la forme *komputilo* (ordinateur), voir le chapitre que lui consacre Lo Jacomo (1981: 175-186).

<sup>34</sup> Selon le corpus Internet de 2004 utilisé par Sten Johansson (2015), on relève 761 occurrences de *inaŭguro* contre 5 occurrences pour *inaŭguracio*.

dynamique du procès. De même, la forme *komentario* "commentaire" reprend le latin *commentarius* (comportant un suffixe *-arius*), mais se trouve simplifiée en *komento*<sup>35</sup>. Toutefois, l'évolution n'est pas toujours achevée, comme le montre par exemple la forme *aŭtentika* "authentique" (< lat. *authenticus* < gr. *αὐθεντικός*). Celle-ci reste toujours utilisée, même si la forme réduite sans le pseudo-suffixe *-ika* (à savoir : *aŭtenta*) gagne également en fréquence d'usage<sup>36</sup>.

Ces deux tendances reflètent l'opposition que nous avons déjà entrevue dans l'histoire des langues construites, à savoir l'opposition entre langues schématiques et langues naturalistes. On peut retrouver cette opposition à l'intérieur même de l'espéranto, avec une tendance schématique (*inaŭguro*, *komento*, *aŭtenta*...) et une tendance naturaliste (*inaŭguracio*, *komentario*, *aŭtentika*...). Ces deux perspectives d'évolution sont discutées et argumentées, par exemple chez Waringhien (1959, 1980) en faveur d'emprunts plus importants aux langues romanes, ou au contraire en faveur d'une plus grande cohérence interne de la langue chez Piron (1987, 1994)<sup>37</sup>.

Il convient de noter que ces possibilités d'évolution étaient déjà indiquées par la Règle 15 du Fundamento, que nous avons précédemment citée. Revenons un instant sur la fin de la formulation de cette règle : « (...) Mais quand, dans une catégorie, plusieurs mots différents dérivent de la même racine, il vaut mieux n'employer que le mot fondamental, sans altération, et former les autres d'après les règles de la langue internationale ». Le choix des termes n'est pas ici innocent : « il vaut mieux... ». Il ne s'agit donc pas d'une obligation ou d'une interdiction, mais d'une piste suggérée (et recommandée) de la part de Zamenhof. Cela montre donc là aussi la flexibilité que Zamenhof souhaitait introduire dans son projet de langue, afin de faciliter le passage du stade de projet à celui d'une langue vivante, apte à évoluer de façon indépendante, et semblable en cela aux langues naturelles.

## 8. Conclusion

Dans le domaine lexical, l'espéranto évolue de façon semblable à ce qui peut s'observer dans les langues naturelles. La seule précaution qu'il convient de prendre ici en ligne de compte est l'existence de deux phases dans la constitution et l'évolution de la langue. Dans une première phase qui concerne la constitution du projet de langue, l'enjeu est de rendre possible un développement ultérieur de la langue, en laissant ouverte différentes potentialités d'évolution. Dans une seconde phase, qui concerne la vie

---

<sup>35</sup> La forme *komentario* est la plus fréquente dans Tekstaro pour la période 1887-1940 (88% des occurrences pour la forme *komentario*), alors que la tendance s'inverse si on considère le corpus Internet 2004 de Johansson (74,6% des occurrences correspondent à la forme *komento*).

<sup>36</sup> À la différence de *aŭtentika*, la forme *aŭtenta* n'apparaît pas dans le corpus Tekstaro pour la période 1887-1940 ; en revanche, elles sont toutes deux aussi fréquentes dans le corpus Internet 2004 de Johansson (223 tokens de *aŭtenta* / 225 tokens de *aŭtentika*). Si cette tendance se confirme, *aŭtenta* pourrait donc devenir la forme la plus fréquente à l'avenir.

<sup>37</sup> Plusieurs exemples que nous mentionnons sont assez exemplaires de ce débat, et se trouvent ainsi commentés par Lo Jacomo (1981) et Gledhill (1998: 116-121). Ce dernier évalue également les tendances évolutives à l'œuvre dans le lexique de l'espéranto, à travers une étude de corpus (voir en particulier la section 4.1. *Lexical paradigms*).

sémiologique du langage, la langue évolue par l'usage même qui en est fait par ses locuteurs, indépendamment de ce qui a pu être son origine.

De fait, si l'on fait un instant abstraction de l'origine historique de l'espéranto, l'espéranto présente des caractéristiques similaires à celles des langues naturelles. Il est donc normal que cette langue puisse évoluer en diachronie, comme cela peut être observé pour d'autres langues. Comme le dit Ilona Koutny : « D'après son usage actuel, l'espéranto fonctionne exactement comme les langues ethniques, il est capable d'exprimer les pensées et les sentiments de ses locuteurs, il se développe au sein de la communauté qui le parle (...) Certains considèrent que l'espéranto est une langue naturelle. Haitao Liu (2011) a prouvé au moyen d'une analyse quantitative que, structurellement, l'espéranto était une langue tout à fait normale. En tout cas, *langues naturelles* et *langues planifiées* ne sont pas deux catégories dichotomiques, mais constituent un continuum » (Koutny 2019: 113).

Pour obtenir ce résultat, il a toutefois fallu au préalable un projet qui soit équilibré, et qui évite les écueils du schématisme et du naturalisme<sup>38</sup>, afin de satisfaire dans le même temps l'objectif d'obtenir une langue aux structures régulières et celui de préserver une forme de proximité avec les langues humaines déjà existantes. Cette position d'équilibre peut sembler instable, mais c'est cette instabilité même qui explique que l'espéranto puisse évoluer et fonctionner. Et c'est d'ailleurs le propre de toute langue humaine que de reposer sur de tels équilibres instables, mêlant contingences et régularités, apports extérieurs et principes analogiques, ce qui conduit finalement toute langue vivante à posséder une dimension historique et diachronique.

---

<sup>38</sup> Sur les écueils auxquels s'exposent les langues philosophiques, voir notamment Sorlin (2012) et Marlaud (2013). Par ailleurs, sur le rôle joué par les autorités linguistiques dans l'évolution de l'espéranto, voir la thèse de Lo Jacomo (1981).

## Bibliography

- Blanke, Detlev. *Interlinguistics and Esperanto Studies: Paths to the scholarly literature*. Esperanto Documents 47A. Rotterdam: UEA, 2004.
- Blanke, Detlev. "Causes of the relative success of Esperanto" *Language Problems and Language Planning (LPLP)* 3 (2009): 251-266.
- Cherpillod, André. *Le suffixe UM*. Couguenard: publié à compte d'auteur, 2011.
- Cherpillod, André. *Konciza Etimologia Vortaro*. Rotterdam: UEA, 2016.
- Couturat, Louis & Leau, Léopold. *Histoire de la langue universelle*. Paris: Hachette, 1903.
- Cowan, John Woldemar. *The Complete Lojban Language*. Fairfax: Logic Language Group Inc, 1997.
- Duc-Goninaz, Michel (coord.). *Nova Plena vortaro de Esperanto*. Paris: SAT, 2020.
- Duc-Goninaz, Michel. "Planification et régularité(s) : quelques réflexions sur l'espéranto" *Travaux du Cercle Linguistique d'Aix-en-Provence* 6 (1988): 79-91.
- Eco, Umberto. *La recherche de la langue parfaite dans la culture européenne*. Paris: Ed. du Seuil, 1994.
- Gledhill, Christopher. *The Grammar of Esperanto. A Corpus-based description*. München: Lincom Europa, 1998.
- Gončarov, Anatolo. *Tiu ĉumalofta sufikso UM*. UEA (édition électronique), 2021.
- Grosjean-Maupin, Emile (red.). *Plena vortaro de Esperanto*. Paris: SAT, 1930.
- Janton, Pierre. *L'Espéranto, Que sais-je ?* Paris: PUF, 1977.
- Johansson, Sten. "Uzi interreton kiel tekstaron por lingvaj esploroj". In *Interlingvo inter lingvoj. Prilingvaj eseoj*, red. Michele Lipari & Humphrey Tonkin, Rotterdam: UEA, 2015.
- Kerckhoffs, Auguste. *Cours complet de Volapük*. Paris: Librairie H. Le Soudier, 1886.
- Koutny, Ilona. "De la planlingvo de Zamenhof ĝis nature disvolviĝanta lingvo: etapo de la evoluo de Esperanto". In *Ludwik Zamenhof okaze de la centa datreveno, Conferenze 141*, Accademia Polacca delle scienze / Biblioteca e Centro di studi a Roma, 2018.
- Koutny, Ilona. "Caractérisation typologique de l'espéranto comme langue naturelle" *Cahiers de l'ILSL* 61 (2019): 111-137.
- Libert, Alan Reed. *Mixed artificial languages*. München: Lincom Europa, 2003.
- Lo Jacomo, François. *Liberté ou autorité dans l'évolution de l'espéranto*. Thèse de doctorat. Université Paris V, 1981.
- Marlaud, Sarah. "Les langues artificielles sont-elles des langues ? Etude contrastive de l'espéranto et de la caractéristique universelle" *Syntaxe & Sémantique* 14 (2013): 85-117.
- Martinet, André. *Eléments de linguistique générale*. Paris: Colin, 1960.

*Investigationes Linguisticae, vol. XLVII*

- Mattos, Geraldo. *La deveno de Esperanto*. Chapecó: Fonto, 1987.
- Monnerot-Dumaine, Marcel. *Précis d'interlinguistique générale et spéciale*. Paris: Librairie Maloine, 1960.
- Pabst, Bernhard. *Berlina Komentario pri la Fundamento de Esperanto*. Berlin: Esperanto-Akademio, 2015.
- Piron, Claude. *La bona lingvo*. Vienne: Pro Esperanto, 1987.
- Piron, Claude. *Le défi des langues. Du gâchis au bon sens*. Paris: L'Harmattan, 1994.
- Planchon, Philippe. "Diakronio kaj evolu-fazoj de la leksiko kaj de la vort-farado en Esperanto" *Internacia Pedagogia Revuo* 4/2021 (2021): 39-49.
- Saussure, Ferdinand (de). *Cours de linguistique générale*. Paris: Ed. Payot, 1995 [1916].
- Slaughter, Mary M. *Universal languages and scientific taxonomy in the seventeenth century*. Cambridge University Press, 1982.
- Sorlin, Sandrine. *Langue et autorité : de l'ordre linguistique à la force dialogique*. Presses Universitaires de Rennes, 2012.
- Stria, Ida. *Towards a linguistic worldview for artificial languages*. Thèse de doctorat, Université Adam-Mickiewicz de Poznań, 2015.
- Van Oostendorp, Marc. "Syllable structure in Esperanto as an instantiation of universal phonology" *Esperantologio / Esperanto Studies* 1 (1999), 52–80.
- Waringhien, Gaston. *Lingvo kaj vivo*. Anvers / La Laguna: Ed. Stafeto, 1959.
- Waringhien, Gaston (red.). *Plena Ilustrita Vortaro*. Paris: SAT, 1970.
- Waringhien, Gaston. *1887 kaj la sekvo...* Anvers / La Laguna: Ed. Stafeto, 1980.
- Wells, John. *Lingvistikaj aspektoj de Esperanto*. Rotterdam: UEA, 1989.
- Wilkins, John. *An Essay towards a Real Character and a Philosophical Language*. Londres: Royal Society, 1668.
- Zamenhof, Ludwig Lejzer. *Международный языкъ. Предисловіе и полный учебникъ 'Langue internationale. Avant-propos et manuel complet' (Unua Libro)*. Varsovie: Ĥaim Kelter, 1887.
- Zamenhof, Ludwig Lejzer. *Dua Libro de l' lingvo internacia*. Varsovie: Ĥaim Kelter, 1888.
- Zamenhof, Ludwig Lejzer. *Fundamento de Esperanto*. Pise: Edistudio, 1991 [1905].
- ⇒ Corpus *Tekstaro* (Esperantic Studies Foundation – Bertilo Wennergren) : <https://tekstaro.com>