



Cahier des charges pour une Interlangue

Dans le genre divertissant, vous avez pu lire récemment "La débabélisation" et "L'Europanto". Plus sérieusement, coexistent depuis de nombreuses années le Volapük (qui est resté très confidentiel) et l'Espéranto (qui a un certain succès d'estime, ayant séduit quelque 0,05% de la population mondiale, soit 1,5 fois l'équivalent de la population de la ville de Paris).

Pourquoi ce succès mitigé ? Essayons de définir ce qu'il faut exiger d'une interlangue à notre époque, et ce qu'il ne faut pas en attendre.

***Avis aux linguistes** : cet article a été rédigé par un amateur, donc il se peut qu'il contienne quelques inexactitudes. L'auteur leur saurait donc gré de ne pas s'en formaliser, mais plutôt de l'en informer, afin que les rectificatifs nécessaires puissent être fait dans le prochain numéro.*

État des lieux

L'espéranto – pour ne parler que de lui – date maintenant de plus d'un siècle (de 1887 pour être précis). Son créateur, le docteur Zamenhof, avait pour objectif¹ de créer un langage :

- dépourvu d'irrégularités, tant dans l'orthographe que dans la prononciation et la grammaire. Objectif atteint.
- en faire une langue vivante. On peut estimer que cet objectif a également été atteint, puisque parlée par plusieurs millions de personnes, et s'enrichissant avec le temps.
- en faire une langue susceptible de se substituer, au moins partiellement, aux langues nationales. Là, c'est l'échec complet : l'espéranto n'est la langue maternelle de personne. De plus, la population espérantiste est actuellement quasiment en croissance zéro. Un autre objectif de l'espéranto a cependant été relativement atteint : disposer d'une langue littéraire (il existe plusieurs dizaines de milliers d'ouvrages traduits en espéranto).

De plus, le Dr Zamenhof étant polonais et du XIX^{ème} siècle, il a raisonné uniquement "monde occidental" et défini des règles inspirées de sa langue natale (pluriels en -oj, etc.), ce qui favorise les habitants de l'Europe de l'est² et peut être considéré comme un inconvénient pour ceux qui ont des habitudes d'élocution très différentes.

Maintenant, nous sommes en 1999. Il y a l'informatique, Internet... Il faut raisonner mondial. À partir de là, quelles sont les règles à définir – et à respecter – pour créer une interlangue susceptible d'être réellement utilisée ? Elles sont simples :

- la première règle, dont dérivent toutes les autres, c'est qu'une interlangue doit être très simple à apprendre et à pratiquer. Donc :
- comme l'espéranto, cette interlangue doit être totalement dépourvue d'irrégularités : prononciation, orthographe, grammaire ; 0 exception est une règle incontournable ;
- elle doit être simplifiée à l'extrême, quitte à user et abuser des préfixes et suffixes pour préciser le sens des mots ;
- elle doit utiliser un alphabet restreint et universellement connu ;

¹ Logique, le Dr Zamenhof était oculiste de métier.

² Officiellement, il a composé "pan-européen", mais j'ai toujours trouvé que l'espéranto (que je n'ai jamais pratiqué, seulement vu et un peu entendu) ressemblait furieusement à un mélange de polonais et de roumain...

- elle doit être aisément prononçable et compréhensible par tous – y compris les ordinateurs en reconnaissance vocale –, quelles que soient les habitudes de prononciation ;
- enfin, il s'agit d'une langue d'échange, qui n'a pas à se substituer aux langues et idiomes nationaux et régionaux : elle peut donc être pauvre, tant que les idées passent bien.

Détaillons ces contraintes une par une.

L'alphabet

Il faut évidemment, pour des raisons de compatibilité, s'en tenir aux 26 lettres de l'alphabet latin, ou même à un sous ensemble de celui-ci, qui sont utilisées partout - même en Chine. Ceci se passe de commentaires.

La prononciation

Note : j'ai utilisé ici l'alphabet phonétique pour mettre en valeur les différences de prononciation. Pour les sons utilisés en France – et quelques autres de nos voisins –, vous pouvez consulter l'annexe en fin d'article. Pour les autres, ils sont montrés à titre indicatif.

Ici, c'est beaucoup plus compliqué (il faut donc simplifier). En effet, les habitudes ne sont pas les mêmes partout, loin de là :

- Certains sons sont très utilisés ici, mais pas là. Par exemple, il existe plusieurs façons de prononcer le R :
 - le R palatin français ;
 - les R roulés (accents africains et russe, l'espagnol...)
 - les R gutturaux (allemand, néerlandais) ;

...mais l'alphabet phonétique international en définit au moins une dizaine (ne m'en demandez pas la prononciation), qui s'écrivent : **R**, **ʁ**, **ɹ**, **ɻ**, **ɽ**, **ɾ**, **ɿ**, **ʀ**, ou encore **ʁ̥**.

...mais les chinois ne font aucune différence entre le L et le R (vous pouvez dire *honorable*, *honorable* ou même *honorabre*, pour eux c'est pareil), ni les arabes entre le B et le P (vous pouvez dire *papa* ou *baba*, pour eux c'est *baba* et ça veut dire *papa*), le nom propre "Pakistan" étant une exception qui a nécessité un caractère spécial rien que pour lui.

- Nous utilisons couramment des sons dont nous ne soupçonnons même pas l'existence. Par exemple, en français, le son **ɲ**, qu'on trouve dans *agneau* (**ɑɲo**) ou *baigner* (**bɛɲe**), et qui n'est pas la diphtongue **ny**. Ou le son **ŋ**, comme dans *camping* (**kɑ̃piŋ**) ou *ping-pong* (**piŋ-poŋ**), qui n'est pas le son **g**.
- Il existe dans la plupart des langues des subtilités de prononciations qui sont plus ou moins importantes. Par exemple, en français, qui se soucie de la différence entre le **a** de *patte* (**pat**) et celui de *pâte* (**pat**)³ Dans un même ordre d'idées, la distinction entre **in** (serein : **sɛrɛ̃**) et **un** (embrun : **ɑ̃brœ̃**) tend à tomber en désuétude. Par contre, en néerlandais⁴, il ne faut pas confondre *boom* (**bom**) (arbre) et *bom* (**bɔm**) (bombe), sinon les jardiniers seraient tous des terroristes ! En effet, dans cette langue, les voyelles longues et les voyelles courtes sont fondamentalement différentes (en = et, een = un ; zon = soleil, zoon = fils, etc.).
- Au sein même d'un même pays il existe des différences de prononciations : en France, wagon se prononce **wagɔ̃** (ouagon) dans le Nord (comme les anglais, les belges et les hollandais), **vagɔ̃** (vagon) à Paris (comme les allemands)⁵. En Belgique, le **g** est tantôt un g

³ Personnellement, avant d'étudier l'alphabet phonétique j'ignorais jusqu'à l'existence de cette nuance !

⁴ J'utilise fréquemment des références au néerlandais, parce que suis certain que très peu d'entre vous le connaissent, et qu'une référence à une langue non connue peut être plus parlante (vous ne risquez pas de dire : c'est normal).

⁵ À ce propos, Napoléon s'est fait battre à Waterloo (Ouaterlo), et pas à 'Vaterlo'; près de Lille, Wattrelos se prononce

(g), tantôt un h aspiré (h), tantôt un r guttural (χ), selon qu'on est (peut-être dans le désordre) à Gand, Anvers ou Ostende.

- Et, bien sûr, certains pays emploient des sons que d'autres sont même incapables de prononcer correctement : les différents coups de glotte (ʔ, ʕ, ʙ, ʔ̥, ʕ̥) ou le clic bilabial (ɸ) pour ne citer que ceux-là...

Comment s'en sortir ? Réponse ci-dessous.

Associer alphabet et prononciation

Bien sûr, chacun sait que dans la plupart des langues alphabétiques, il n'y a pas bijection entre une lettre et sa prononciation. Exemples :

- En français, la lettre **c** a trois prononciations différentes : **k** dans *tracas*, **s** dans *cinéma*, **ʒ** dans *seconde*.
- Inversement, le son **s** peut être représenté par **s** (*singe*), **ss** (*chasse*), **c** (*civet*), **x** (*Bruxelles*⁶), ou même **t** (*potion*)...
- Nous changeons parfois la prononciation de certaines lettres pour de simples raisons de facilité, mais cela ne modifie pas le sens du mot : le **o** de *cote* (kɔt) et celui de *côté* (kote), le **eu** de *heure* (œr) et celui de *heureux* (œrø), etc.
- Entre pays différents, ce n'est pas mieux : si en France la lettre **j** se prononce ʒ (c'est-à-dire... j), les anglais prononcent dʒ (*dj*), les néerlandais j (y), les espagnols x (r roulé)...
- Il y a aussi le problème de l'accent tonique. Si en français et dans d'autres langues le fait de mal le placer n'est qu'un défaut de prononciation, ce n'est pas vrai partout. Dans certaines langues le sens d'un mot peut dépendre de cet accent tonique. Exemple près de chez nous : en espagnol, *el papa* signifie *le papa*, mais *el papa* désigne *le pape* ! Dans les langues asiatiques, c'est pire : on peut passer du compliment à l'injure rien qu'en se trompant sur l'emplacement de l'accent tonique (ce qui complique la langue écrite : le même son s'écrit différemment selon son degré de "tonicité") ! Et voilà pourquoi l'arabe (entre autres) est plein de points et de virgules partout autour des lettres (ce qui nécessite en Unicode plusieurs centaines de symboles !).

Bref, impossible de dire une lettre = un son (c'est pourtant ce que fait l'alphabet phonétique, mais au prix d'un alphabet d'environ 120 lettres, surtout que si nous sommes incapables (sauf peut-être avec beaucoup d'entraînement) de parler avec des clics et des coups de glotte, tout près de nous les espagnols ont du mal à prononcer notre **j**, les italiens à dire **u** et pas **ou**...

Pour que notre interlangue soit facile pour tout le monde, une seule solution : le nivellement par le bas. Il faut qu'un son ne puisse être transcrit qu'avec une seule lettre possible, mais permettre des prononciations différentes (mais cependant proches) pour une lettre donnée. Ceci permet une multitude d'"accents" pour une seule écriture et limite au maximum les problèmes de compréhension entre nationalités différentes.

Bien sûr, il faudra "sacrifier" certaines distinctions et considérer des sons que nous sommes habitués à différencier comme équivalents. Selon toute probabilité, ce devrait être le cas du **u** (qui s'écrit phonétiquement y) avec le **ou** (qui s'écrit phonétiquement u), du **a** (a ou ɑ) avec le **an** (ã ou ɞ̃), du **o** (o ou ɔ) avec le **on** (õ ou ɔ̃), etc.

Quant à l'accent tonique, c'est une intonation, pas un son. Il ne peut donc dans notre interlangue avoir aucune influence sur le sens des mots et n'intervient ainsi que comme liberté de prononciation.

Quatrelo et pas 'Vatrelosse' comme l'écorchent si bien nos journalistes parisiens.

⁶ *Qui se prononce Brussel et pas Bruksel. (Avis aux parisiens).*

Grammaire et syntaxe

En simplifiant au maximum, nous devons définir des règles pour les catégories suivantes :

- accords de genre (masculin, féminin, singulier, pluriel, ...) ;
- conjugaison des verbes ;
- dérivations de sens des mots ;
- articles, conjonctions, etc.

Mais il faut respecter, pour que l'interlangue soit facile d'apprentissage, une règle d'or : un phonème correspondant à un préfixe, suffixe ou conjonction ne peut être réutilisé pour un autre usage ; et inversement, il ne faut pas permettre à deux phonèmes différents d'avoir le même usage. C'est par exemple une des principales difficultés de celui qui apprend le néerlandais :

- Selon les mots, le pluriel est soit en **-s**, soit en **-en** (mais ces deux suffixes sont généralement exclusifs, seuls quelques mots pouvant être accordés des deux façons). Résultat : la simple liste des règles d'accord du pluriel en néerlandais occuperait une page entière de La Lettre !⁷
- De plus, le suffixe **-en** concerne non seulement le pluriel, mais est également la marque de l'infinitif : *werken* peut se traduire, soit par *travailler*, soit par *travaux*. Seul le contexte permet de lever l'ambiguïté.

Globalement, le meilleur moyen de bien construire une langue synthétique semble être de définir les phonèmes, puis de leur attribuer un usage. Ceci devrait éviter le risque d'homophonies (par exemple trois mots qui se suivent ressemblant à une autre suite, de deux mots, signifiant tout autre chose). Bien sûr, les homophonies sont couramment utilisées en français humoristique, sous l'appellation de calembours ou jeux de mots, mais l'interlangue n'a pas besoin de cela, compte tenu de son usage.

Pourquoi utiliser un tel principe ? Parce que moins il y a de mots à apprendre, plus l'apprentissage est rapide. Un exemple : si à l'aide de 200 radicaux et 10 préfixes et suffixes, on peut obtenir par combinaison 4000 mots différents, combien de temps faudra-t-il pour apprendre ces 4000 mots ? Le temps d'apprendre les 210 phonèmes de base...

Bien que plusieurs milliers de syllabes distinctes puissent être définies, il faudra toujours faire en sorte que l'association de deux syllabes A et B ne donne pas le même résultat que les syllabes C et D (par exemple éviter que les syllabes *crap* et *ul* puissent se suivre si les syllabes *cra* et *pul* le peuvent également). Si cela pose problème, il faudrait peut-être prévoir un séparateur qui puisse lever toute ambiguïté (un son réservé à cet usage).

Et pour définir les règles, comparons ce qui se fait dans différentes langues et choisissons la méthode la plus simple.

Les accords de genre

Pour le pluriel, pas de difficulté majeure : on choisit un phonème qui sera exclusivement réservé à cet usage, et qui se mettra au début ou à la fin du mot pour lui donner la marque du pluriel. Ceci est plus simple que la méthode française (le **-s** final étant souvent muet), qui nécessite l'ajout d'un article accordé pour la compréhension orale : pour dire qu'il y a plus d'un chien, nous sommes obligés de dire *des chiens*, alors qu'il suffit aux anglais de dire *dogs*.

Pour le genre, c'est un peu plus complexe. En français, le neutre se confond avec le masculin. En néerlandais et en allemand, n'importe quel nom peut être de l'un des trois genres⁸. En anglais, tout ce qui n'est pas sexué est neutre. Il semble qu'il n'y ait pas photo sur ce point.

⁷ Dans ce domaine, nous sommes assez favorisés avec le français : le **-s** est de rigueur, sauf pour les mots en **-al** ou **-ail** qui deviennent **-aux**, avec quelques exceptions comme **naval**, les mots finissant au singulier par un **s** ou un **x** qui ne changent pas, et bien sûr nos sept mots en **-ou** qui prennent un **x** (il y a maintenant un huitième mot : **ripou-x**).

⁸ Et comme en néerlandais l'article **de** correspond aussi bien au masculin qu'au féminin, je n'ai pas réussi à savoir – ni en me renseignant ni même avec un dictionnaire ! – si **de tabel** (la table) est masculin ou féminin. Tout ce que je sais, c'est que ce mot n'est pas du genre neutre (sinon ce serait **het tabel**). Il semble que les néerlandais eux-mêmes n'en sachent rien.

Si pour le singulier c'est simple (un nom est soit, masculin, soit féminin, soit neutre... quoique reste posé le cas des hermaphrodites), ça l'est moins pour le pluriel : si on désigne deux ou plusieurs "personnes", elles peuvent être toutes du même genre, être de deux ou plusieurs genres différents, et même on peut ignorer la composition du groupe, ou celle-ci peut être sans importance... Faut-il définir une demi-douzaine d'accords de genre pour le pluriel ? Ceci reste à définir.

Conjugaison des verbes

Pour les "personnes", il existe deux écoles principales : faire porter l'accord sur le verbe (comme en latin : *habeam, habeas, habeat, habemus...* – pour *j'ai, tu as, il a, nous avons* – lequel se passe parfaitement des pronoms personnels), ou sur le pronom (comme en anglais : *I cut, you cut, we cut* – *je coupe, tu coupes, nous coupons*). En français, nous cumulons les deux, ce qui est redondant et ne facilite pas la vie des écoliers.

Sachant que le pronom personnel, placé juste avant le verbe, peut être considéré comme un préfixe, et qu'il peut servir ailleurs, je pencherais pour cette dernière solution, réservant l'usage d'un vrai préfixe ou suffixe pour définir le temps.

Dérivations de sens des mots

Dériver le sens d'un mot, c'est créer une famille de verbes, substantifs, adjectifs, adverbes... à partir d'un radical de base (par exemple, en français, *vert* nous donne *verdir, verdure, vertement, verdâtre...*).

Il suffit donc de définir, d'une part les phonèmes des substantifs de base, et d'autre part ceux des préfixes et suffixes donnant les variations, et nous obtenons l'essentiel de l'interlangue.

Par exemple, *j'ai mangé* pourrait se dire (avec des phonèmes plus réduits, bien sûr) *moi mange-temps-derrière*.

Bien sûr, il pourrait arriver qu'à force d'assembler préfixes et suffixes on se retrouve avec des mots pleins de syllabes. Donc, dans certains cas extrêmes concernant des notions d'usage courant, il faudrait prévoir des phonèmes de remplacement, un mot de base très long pouvant avoir un synonyme court.

Articles, conjonctions, etc.

Les articles sont – du moins dans les langues indo-européennes – des syllabes placées devant les substantifs pour leur apporter un complément de précision. Si l'article indéfini peut être omis, comme on l'a vu plus haut, il semble que les articles définis et démonstratifs puissent être définis comme des adverbes (les substantifs étant déjà accordés, les articles peuvent être invariables).

Quand aux conjonctions, prépositions, etc., définissons des syllabes de base qui puissent s'associer avec les différents préfixes et suffixes de dérivation ou d'accord, et on devrait avec relativement peu de mots obtenir un vocabulaire assez complet.

Le problème des noms propres

Il reste un dernier cas à traiter : tous les mots appartenant à la première partie du Petit Larousse sont à définir, mais que faisons-nous de ceux de la seconde partie ? Par définition, un nom propre ne se traduit pas (il y a bien quelques exceptions avec les noms très connus : par exemple *Deutschland* se dit en français *Allemagne* et en anglais *Germany...*).

Personnellement, je ne vois qu'une seule solution : définir deux phonèmes signifiant "début de nom propre" et "fin de nom propre" permettant de délimiter un groupe de phonèmes ne suivant pas les règles de l'interlangue (et où la prononciation peut être beaucoup plus stricte), tout comme dans un programme source on délimite les chaînes de caractères par des apostrophes ou des guillemets pour les différencier du reste du code.

Et si un nom propre contient lui-même le phonème de fin de nom propre ? Bonne question...

Et UNL ?

L'UNU (United Nations University) a lancé en 1996 un ambitieux programme de création d'un langage universel : UNL (Universal Network Language), mais dont le but est uniquement de créer une interface standard pour le langage écrit (quelque 300 000 "mots universels" ont été définis) :

- Vous tapez votre texte en français, l'encodeur le stocke en UNL.
- Si un Japonais veut lire votre texte, son décodeur lui affichera du japonais.

L'ambition est qu'UNL s'incorpore à HTML ou XML : fini les sites Web à 99% anglophones !

Par ailleurs, deux remarques importantes :

- Si les décodeurs sont déjà au point, les encodeurs posent beaucoup plus de problèmes ;
- La codification UNL n'est pas linéaire, et n'a rien à voir avec une langue orale (un document UNL est d'ailleurs encore plus illisible qu'un programme en C écrit par un bidouilleur fou).

Bref, une interlangue reste nécessaire pour la communication non écrite, mais sa syntaxe ne doit poser aucune difficulté pour un encodeur UNL.

Pour ceux qui veulent en savoir plus sur UNL :

- Sur le Web : unl.ias.unu.edu
- Auprès d'ADELI : juste un peu de patience, un article est en préparation pour le n° 35 de La Lettre.

Conclusion

En résumé, nous pouvons créer une véritable interlangue en trois étapes :

- 1) Définir l'alphabet et les sons associés ;
- 2) Inventorier les phonèmes possibles et les possibilités d'assemblage (en tenant compte d'UNL dans la mesure du possible) ;
- 3) Créer le vocabulaire.

...sans oublier de :

- 4) lui donner un nom ;
- 5) s'en servir.

Par ailleurs, je regrette, dans cette analyse, de n'avoir pu utiliser aucune notion de langues non indo-européennes (chinois, japonais, coréen, dialecte africain, etc.), ce qui la rend forcément incomplète et favorise l'indo-européen. Je souhaite que quelqu'un ayant de telles notions (il y en a à ADELI) apporte des idées complémentaires.

Et si on s'y mettait ?...▲

Jean-Luc Blary
e-mail : jlblary@nordnet.fr

Annexe 1 : extrait de l'alphabet phonétique international

Voyelles orales		Voyelles nasales		Semi-voyelles	
<i>il, habit, dîner</i>	i	<i>intérêt, pain, sein</i>	ẽ	<i>yoyo, caillou</i>	j
<i>thé, dé</i>	e	<i>alun, parfum</i>	œ̃	<i>huile, lui, huissier</i>	ɥ
<i>être, dais, procès</i>	ɛ	<i>entrer, blanc</i>	ã	<i>oui, wagon</i>	w
<i>avoir, Paris, patte</i>	a	<i>ondée, bon, honte</i>	õ		
<i>âne, pâte, mât</i>	ɑ	Consonnes			
<i>or, robe, maure</i>	ɔ	<i>pas, dépasser, cap</i>	p	<i>zone, raison, gaz</i>	z
<i>dos, chevaux</i>	o	<i>tu, étaler, lutte</i>	t	<i>cheval, mâcher, match</i>	ʃ
<i>ouvrir, couvert, loup</i>	u	<i>caste, képi, que</i>	k	<i>jambe, âgé, page</i>	ʒ
<i>user, tu, sûr</i>	y	<i>beau, abîmer, club</i>	b	<i>large, mollesse, mal</i>	l
<i>cœur, peur, neuf</i>	œ	<i>dur, broder, bled</i>	d	<i>rude, mari, ouvrir</i>	ʀ
<i>feu, jeu, peu</i>	ø	<i>gare, vague, zigzag</i>	g	<i>maison, amener, blême</i>	m
<i>le, premier</i>	ə	<i>fou, affreux, chef</i>	f	<i>nourrir, fanal, dolmen</i>	n
Divers		<i>vite, ouvrir</i>	v	<i>agneau, baigner</i>	ɲ
<i>h "aspiré" muet (haricot)</i>	*	<i>souffler, chasse, hélas</i>	s	<i>camping</i>	ŋ
Autres sons (peu ou pas utilisés en France)					
<i>h prononcé (hé, hé !)</i>	h	<i>ch chuinté (ich allemand)</i>	ç	<i>u anglais (duck)</i>	ʌ
<i>r espagnol (jota, Julio)</i>	x	<i>r soufflé (ach allemand)</i>	χ	<i>th anglais (thin)</i>	θ
				<i>th anglais (that)</i>	ð

Annexe 2 : parler n'est pas souffler

Selon qu'on parle haut ou bas (c'est-à-dire qu'on fait ou non vibrer ses cordes vocales), on ne prononce pas les mêmes sons. Voici une petite comparaison (en alphabet phonétique) :

haut	b	d	v	g	ʒ	z	ʀ
bas	p	t	f	k	ʃ	s	χ