

# Une perspective orwellienne sur les tentatives d'amélioration de la communication : Du Basic English au web sémantique

Roger BAUTIER\*

**Résumé :** L'intérêt manifesté par George Orwell à l'égard du Basic English, qu'il a approuvé puis condamné, doit être replacé dans la longue histoire de la rationalisation de la communication. Une histoire qui se poursuit actuellement par l'essor du web sémantique, dont les

développements peuvent susciter, tout autant, des enthousiasmes ou des critiques.

**Mots-clés :** Orwell, Basic English, rationalisation, web sémantique, ontologies

\*\*\*

*An Orwellian perspective on attempts to improve communication:  
From Basic English to semantic web*

**Abstract:** The interest shown by George Orwell in Basic English, which he approved, and then sentenced, must be seen in the long history of the rationalization of communication. A story that is ongoing by the rise of semantic web, including developments that can generate, just as much, enthusiasm or

criticism.

**Keywords:** Orwell, Basic English, rationalization, semantic web, ontologies

\*\*\*

---

\*Professeur en Sciences de l'Information et de la Communication, Université Paris 13, Villetaneuse, LABSIC, roger.bautier@laposte.net.

Si l'on considère l'élaboration de tout dispositif de médiation comme une tentative de remédier à des défauts de communication, il est remarquable qu'une telle entreprise soit rarement exempte d'ambiguïté. Un bon exemple en est sans doute celui de la réflexion menée par George Orwell sur le discours politique et le discours journalistique, dont on ne perçoit le plus souvent qu'un aspect. En effet, il est bien connu que, dans son roman d'anticipation intitulé *1984*, George Orwell conçoit la standardisation linguistique propre à la Novlangue (Newspeak) comme un moyen d'empêcher l'expression de toute pensée hérétique. On considère en général qu'elle constitue notamment une critique du Basic English qui avait été proposé par Charles Ogden à la fin des années 20 pour servir de langue artificielle internationale qui permettrait un contrôle du sens des énoncés. En revanche, ce que l'on sait moins, c'est que le projet formulé par Ogden avait intéressé Orwell, qui, lui aussi préoccupé par les problèmes soulevés par l'usage de la langue, avait envisagé, quelques années avant la publication de *1984*, un certain nombre de précautions à prendre afin d'améliorer la communication en général et, plus particulièrement, celle qui passait alors par la presse et la radio.

Il paraît donc nécessaire de se pencher sur les particularités de la perspective propre à Orwell, dont l'instabilité, au moins apparente, interpelle nécessairement le lecteur de son œuvre. C'est la reconnaissance de sa complexité qui permet ensuite de replacer sa démarche dans un vaste ensemble de préoccupations à l'égard de cette amélioration, comme celles qui se sont manifestées chez les idéologues de la Révolution française ou chez ceux de la Troisième République. Cette reconnaissance permet également d'examiner la référence fréquente faite à Orwell dans les considérations sur le fonctionnement des discours de propagande au cours de la seconde moitié du XXe siècle. C'est elle, enfin, qui invite et introduit à l'analyse du renouvellement de la problématique initiale, renouvellement qui a été entraîné par le développement des nouvelles technologies de l'information et de la communication et, notamment, par les tentatives récentes visant à constituer un web dit « sémantique ».

### **Le Basic English et les jugements d'Orwell**

En 1944, dans l'hebdomadaire *Tribune*, Orwell, qui est un écrivain et journaliste déjà célèbre, publie une chronique dans laquelle il critique ainsi les hommes politiques et les journalistes (Orwell, 1944) : « One argument for Basic English is that by existing side by side with Standard English it can act as a sort of corrective to the oratory of statesmen and publicists. High-sounding phrases, when translated into Basic, are often deflated in a surprising way. For example, I presented to a Basic expert the sentence 'He little knew the fate that lay in store for him' - to be told that in Basic this would become 'He was far from certain what was going to happen'. It sounds decidedly less impressive, but it means the same. In Basic, I am told, you

cannot make a meaningless statement without its being apparent that it is meaningless - which is quite enough to explain why so many schoolmasters, editors, politicians and literary critics object to it. » Le Basic English auquel Orwell fait référence avait été présenté par Ogden dans un ouvrage publié en 1930. L'élaboration d'une telle langue réduite correspond à la prise de conscience de la nécessité d'une « orthologie » : le Basic English doit favoriser l'exercice de la fonction référentielle du langage en donnant une place prépondérante à un emploi « correct » des mots. C'est pourquoi les langues existantes peuvent apparaître comme des outils médiocres de médiation entre les humains, voire comme des obstacles qui ont peut-être été responsables de calamités (Ogden, 1930) : « Before the Great War, it was clear to most persons with a knowledge of history and an interest in international organization that one of the chief needs of Europe was fifty more dead languages. » Le travail d'Ogden aboutit à la proposition d'une langue capable de servir de moyen de communication universel parce que fondée sur l'expression d'un certain nombre de concepts fondamentaux : une langue à la fois restreinte (850 mots, avec très peu de verbes) et très précise (grâce aux définitions adoptées). La démarche qui est la sienne est censée trouver sa validation dans une triple affirmation. Premièrement, selon lui, beaucoup de mots habituellement utilisés peuvent être remplacés par d'autres mots ayant l'avantage d'être « clearly, in some sense, nearer to the facts » ; deuxièmement, beaucoup de choses dont on parle ne sont que des « fictions » qu'il s'agit d'éviter ; troisièmement, il est souhaitable d'éliminer ce qu'il y a d'émotif dans les expressions usuelles.

Le Basic English ne laissera pas Churchill indifférent. Dans un discours prononcé en 1943 devant les étudiants de l'Université Harvard (Churchill, 1943), il rappelle qu'il a demandé au gouvernement britannique d'étudier les potentialités du Basic English et il met en lumière ce qui lui paraît le trait principal de ce moyen de communication : « Here you have a plan. There are others, but here you have a very carefully wrought plan for an international language capable of a very wide transaction of practical business and interchange of ideas. » Il tente même de rallier Roosevelt à sa cause, à travers un échange de lettres (Churchill, Roosevelt, 1944). La réaction de Roosevelt, qui n'est pas hostile, est néanmoins assez ironique : « Incidentally, I wonder what the course of history would have been if in May 1940 you had been able to offer the British people only « blood, work, eye water and face water », which I understand is the best that Basic English can do with five famous words. Seriously, however, we are interested and will look into the matter thoroughly ».

Orwell n'a pas exactement la position de Churchill (la fonction du Basic English comme langue internationale le touche moins), ni celle de Roosevelt (il n'aurait probablement pas défendu la valeur stylistique de la phrase de Churchill : « I have nothing to offer but blood, toil, tears and sweat. »). Cependant, il a échangé avec Ogden à propos des caractéristiques du Basic English et il n'a pas trouvé mauvais

qu'il en soit fait la promotion dans des émissions radiophoniques à destination de l'Inde. Surtout, entre 1944 et 1946, Orwell insiste sur les défauts caractéristiques des discours qu'il lit et qu'il entend, en ayant recours à une normativité qui n'est pas sans rapport avec celle qui est à l'œuvre dans le Basic English. Les textes journalistiques et politiques lui semblent marqués par de regrettables insuffisances (Orwell, 1946) : « The writer either has a meaning and cannot express it, or he inadvertently says something else, or he is almost indifferent as to whether his words mean anything or not. This mixture of vagueness and sheer incompetence is the most marked characteristic of modern English prose, and especially of any kind of political writing. » Il en résulte une analyse très discutable de la production du discours, ainsi que des relations entre l'abstrait et le concret. D'une part, en effet, la signification est considérée comme donnée et c'est à elle que l'on doit permettre de « choisir le mot » qui lui correspond. D'autre part, penser à un objet concret est vu comme autorisant la détermination du mot adéquat, alors que penser à quelque chose d'abstrait est vu comme impliquant la prééminence de mots qui vont, sauf effort conscient pour les en empêcher, obscurcir ou transformer la signification initiale. Il en résulte également une attaque des discours tenus par tous les partis (des conservateurs aux anarchistes) et, en corollaire, un éloge de la simplification de la langue. En simplifiant son expression, on est censé, selon Orwell, pouvoir éviter les « worst follies of orthodoxy », dans la mesure où la simplification constitue à ses yeux une garantie permettant, sinon de ne pas formuler une stupidité, du moins de se rendre compte par soi-même de cette stupidité.

Trois ans plus tard, dans l'ouvrage le plus connu d'Orwell, *Nineteen Eighty-Four. A novel*, la référence à l'orthodoxie est encore présente et la standardisation linguistique caractéristique de la Novlangue qui est la langue officielle d'Oceania, sous la dictature de l'Ingsoc, est présentée de la manière suivante (Orwell, 1949) : « Newspeak was designed not to extend but to diminish the range of thought, and this purpose was indirectly assisted by cutting the choice of words down to a minimum. » Le but est explicité : « The purpose of Newspeak was not only to provide a medium of expression for the world-view and mental habits proper to the devotees of Ingsoc, but to make all other modes of thought impossible. » La limitation de la pensée et l'impossibilité de sortir du cadre de pensée imposé sont considérées comme des conséquences non seulement de la raréfaction du vocabulaire (ainsi que de l'invention de certains mots), mais aussi de la raréfaction des significations : « This was done partly by the invention of new words, but chiefly by eliminating undesirable words and by stripping such words as remained of unorthodox meanings, and so far as possible of all secondary meanings whatever. (...) The word 'free' still existed in Newspeak, but it could only be used in such statements as 'This dog is free from lice' (...). It could not be used in its old sense of 'politically free' or 'intellectually free', since political and intellectual freedom no longer existed even as concepts, and were therefore of necessity nameless. »

Pour la plupart des commentateurs (notamment Courtine, 1984), la Novlangue ainsi définie constitue clairement une critique du Basic English. Ce qui n'empêche pas que l'entreprise d'Ogden a bel et bien intéressé Orwell. Il est possible, certes, d'envisager cette opposition entre les deux moments comme la traduction d'une évolution dans la pensée d'Orwell, qui se serait rendu compte que la réduction de la langue risquait de produire, justement, des effets opposés à ceux qu'elle avait pour fonction déclarée de contrecarrer. Il est possible, également, de l'interpréter comme une contradiction entraînée par une analyse insuffisante des propriétés de la communication s'effectuant par l'usage du langage. Il est possible, enfin, de distinguer radicalement les deux moments, en considérant que la mise en lumière des dangers de la normalisation par la Novlangue est indépendante de tout jugement porté sur le Basic English. Mais, avec le recul de son œuvre dans le temps, ne peut-on pas, tout autant, concevoir, de manière plus dialectique, cette opposition comme une invitation, pour les lecteurs du début du XXI<sup>e</sup> siècle, à reconnaître les tensions incontournables et les nombreuses difficultés auxquelles se heurte nécessairement toute réflexion sur le langage qui se donne pour objectif de concourir à une amélioration de la communication entre les êtres humains ?

### **La rationalisation de la communication**

Le Basic English doit, en effet, être rattaché à toute une réflexion sur la terminologie, réflexion qui n'est elle-même qu'une partie d'un grand projet de rationalisation des langues caractéristique de la première moitié du XX<sup>e</sup> siècle, ainsi que l'a souligné Slodzian (2007). En particulier, il est nécessaire, afin de mieux comprendre son émergence et ses implications, de prendre en considération le terreau qu'a probablement constitué le courant philosophique du positivisme logique. En effet, celui-ci a pu favoriser une vision des langues naturelles qui les ravalait au rang d'obstacles (par leur diversité et leur contingence) au dévoilement d'une vérité universelle, en comparaison, notamment, avec la perfection du langage des mathématiques. Mais il n'est pas difficile de trouver des considérations similaires dans un temps plus ancien.

La Révolution française fournit ainsi des exemples d'une préoccupation pour la rationalisation de la langue, comme chez Condorcet, Talleyrand ou Lakanal. Le privilège à accorder aux sciences est souvent affirmé : Condorcet (1791, 1792) relie la place éminente qui leur est réservée au fait que « la langue en est plus parfaite, que les mêmes mots y expriment plus exactement les mêmes idées ». Ce privilège a une légitimité d'autant plus indiscutable que la conception de la communication qui est prônée par Condorcet implique l'usage d'une parole rationnelle. Parole rationnelle des représentants du peuple, d'abord, qui, étant donné que leur fonction est d'éclairer le peuple et non de l'émouvoir, doivent s'abstenir, sous peine de trahir leur devoir, de toute éloquence « entraînant, passionnée, séductrice » et dont le discours -

imprimé et diffusé largement - sera jugé par des citoyens se constituant en « juges froids et sévères ». Parole rationnelle de l'ensemble des citoyens, ensuite, puisque l'enseignement qui doit être fourni à tous est celui de « l'art de faire des discours écrits » - c'est-à-dire la « véritable rhétorique des modernes » - et, plus précisément, de « l'art d'écrire un mémoire ou un avis avec clarté, avec simplicité, avec méthode ». Un même souci de la communication la plus propre à l'exercice de la raison s'exprime dans les jugements portés au même moment sur les langues qu'il s'agit d'enseigner. Pour Talleyrand (1791), par exemple, il n'est pas seulement nécessaire de faire disparaître les « dialectes corrompus, derniers restes de la féodalité » et de remplacer leur usage par celui de la « langue nationale » qui, seule, permet une communication généralisée ; il faut également ne pas hésiter à procéder au perfectionnement de cette « langue nationale ». Ce perfectionnement passe par un examen critique des mots de la langue et par l'élimination des « significations vagues et indéterminées, si commodes pour l'ignorance et la mauvaise foi ». Le but à atteindre est donc, d'une part, une épuration de la langue destinée à supprimer toute marque de préjugés et à rejeter les formes superflues parce que trompeuses, d'autre part, une stabilisation des signes destinés à transmettre les idées - notamment politiques -, de façon que leur soit attaché « un sens précis et uniforme » qui leur confère un pouvoir de médiation satisfaisant. Là encore, la référence aux sciences est essentielle. En effet, le mode de communication visé correspond à l'application de l'esprit d'analyse et de la méthode des mathématiciens. Chez Lakanal (1794), le lien est explicite : les dogmes politiques nouveaux sont acceptés non par idolâtrie, comme sous l'Ancien Régime, mais par consentement à la rigueur de démonstrations qui ne se différencient aucunement de celles qui ont cours dans les sciences exactes.

On retrouvera, dans la France de la fin du XIXe siècle, la même volonté de rationalisation de la communication, chez les idéologues de la Troisième République soucieux de mettre en place des enseignements qui évacuent la rhétorique, dont il s'agit à nouveau de s'affranchir pour échapper aux perversions d'une autre sorte d'ancien régime et pour établir définitivement la République. Ainsi, Lanson, à l'origine des réformes essentielles des années quatre-vingt concernant les études de lettres, reprend largement le programme défini par le courant rationaliste de la Révolution française (Lanson, 1903). Selon lui, les études littéraires doivent appuyer les études scientifiques. Il faut, selon cette conception, encourager le développement du « sens du vrai » en refusant d'enseigner des mots que les élèves ne peuvent comprendre ; il faut, surtout, condamner le « verbalisme creux » et revendiquer l'usage du « sens propre ». Résolument opposé à un enseignement littéraire dont la fonction semble être de séparer de la foule ceux qui en bénéficient, Lanson insiste sur le fait que « la démocratie demande autre chose ». Elle demande en particulier que l'on soit moins sensible, devant un discours, à la « brillante invention » qui a le défaut de donner « l'habitude de ne pas regarder ce qui est ». L'idéal de la communication se définit alors aisément : c'est celle qui privilégiera « l'exposé exact, ordonné et lumineux, sans éloquence, sans poésie, sans artifice littéraire, d'une

question déterminée, dont la solution dépend d'un choix et d'un examen de faits ». De même, pour Croiset (1903), dans la mesure où la science doit inspirer l'éducation dans une démocratie, il est nécessaire d'apprendre à « ne pas se payer de mots ou d'apparences », à se soumettre à la réalité et à « procéder méthodiquement à l'étude des choses ». Il s'agit là simplement de l'application des méthodes de la science à l'éducation de l'intelligence et, plus largement, à la formation du citoyen.

L'évocation initiale du Basic English par Orwell, d'une part, puis la terrible description qu'il fait de la Novlangue d'Oceania, d'autre part, peuvent donc aisément s'inscrire dans l'histoire fort longue des projets de rationalisation linguistique, mais elles ont la particularité d'attirer l'attention, par l'évolution ou par l'ambiguïté de leur auteur, sur les dangers potentiels de tels projets. On peut remarquer, à ce propos, que la seconde a, jusqu'à la fin du XXe siècle, eu tendance à faire oublier la première, d'autant plus facilement que les critiques du Basic English ont été nombreuses et que la référence à la Novlangue a, par ailleurs, été unanimement utilisée pendant plusieurs décennies pour condamner toutes les « langues de bois ». Qu'il s'agisse de la « langue de bois » qui était censée caractériser les discours communistes ou qu'il s'agisse de la « langue de bois » dont il aurait été possible de repérer les traits dans les médias occidentaux. C'est ainsi que Palmier (1982) souligne ce qui séparerait radicalement les particularités linguistiques de Lénine de celles de Staline. Selon lui, à l'incantation, au faux langage quotidien, à l'appel au pathos qui seraient caractéristiques du style de Staline, répondrait le style de Lénine, emprunt d'une méfiance à l'égard de toute phraséologie, d'un souci constant de la concision allié au rejet de toute emphase, d'une recherche systématique, enfin, des formulations qui refusent la polysémie pour privilégier la précision et l'adéquation du vocabulaire, le recours à l'analyse plutôt que la réalisation d'un effet émotionnel. Du coup, alors même que la Novlangue décrite par Orwell est souvent considérée par les critiques du système soviétique comme la quintessence du langage totalitaire, elle peut, selon Palmier, servir de base pour l'évaluation du langage politique contemporain. Plus largement, elle donnerait des indications précieuses pour l'analyse du langage des communications de masse, notamment pour la mise en évidence de la confusion des mots et des choses, de la traduction du négatif en positif, du contrôle des significations, de l'utilisation de clichés rendant impossible toute opposition, etc., qui seraient spécifiques de ce langage des communications de masse. De même, Sériot (1986) est amené à défendre l'idée que les critiques adressées par les anticommunistes à la propagande communiste et celles formulées par les communistes à l'égard de la propagande anticommuniste manifestent des arguments comparables. Il s'agit, selon lui, presque toujours de mettre en cause le rapport au réel qui est censé fonder le discours étudié et de rechercher, derrière ce discours, la réalité dont il serait possible de parler avec d'autres mots : les « vrais » mots qui diraient la « vérité ».

### **Les espoirs du web sémantique**

À partir du début du XXI<sup>e</sup> siècle, il semble que soient moins présentes les implications politiques qui étaient certes bien visibles dans les propos d'Orwell ou bien dans ceux d'Ogden, et, tout autant, chez ceux qui ont voulu établir la langue de la République française comme chez ceux qui ont bataillé sur l'analyse des discours de propagande. Désormais, la réflexion paraît bien davantage marquée par des préoccupations techniques, du moins apparemment. En même temps, ce qui est frappant, c'est que cette réflexion nouvelle renoue, au moins partiellement, avec l'entreprise illustrée par le Basic English. Les recherches portant sur le web dit « sémantique » en constituent sans doute une bonne illustration, dans la mesure où elles se donnent pour objectif de parvenir à une utilisation perfectionnée des ressources du web par les humains et par les machines. Cet objectif semble, en effet, ne pouvoir être atteint que par la mise au point d'un programme d'interopérabilité entre toutes sortes de données qui, le plus souvent, postule, d'une part, l'indépendance entre les connaissances représentées et les langues dans lesquelles elles sont produites et transmises, d'autre part, leur représentabilité systématique par un formalisme logique. C'est, en tout cas, ce type de programme qui est envisagé par les premiers promoteurs (Berners-Lee, Hendler, Lassila, 2001), à partir de la définition suivante : « The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. »

Le web sémantique va être considéré comme une infrastructure destinée à diminuer la charge des utilisateurs du web : les tâches habituellement effectuées par ces utilisateurs, qui ont à traiter des sources d'information à la fois diverses et hétérogènes, devraient être facilitées par le fait que les machines seraient rendues capables d'accéder aux contenus des ressources existantes et de procéder à des raisonnements à partir de ces contenus. D'où la possibilité d'envisager le web sémantique comme un système de médiation perfectionné reposant sur l'intégration des différentes sources d'information (Laublet, Reynaud, Charlet, 2002). Dès lors, le W3C (World Wide Web Consortium, fondé par Berners-Lee en 1994), diffuse un certain nombre de recommandations concernant les langages susceptibles de participer à la mise au point du web sémantique et propose, à partir de 2004, le standard OWL (Ontology Web Language). Ce dernier repose lui-même sur l'usage de standards plus anciens, notamment XML (eXtensible Markup Language), pour la structuration des documents, et RDF (Resource Description Framework), pour la description des données et des métadonnées sous forme de graphes constitués de triplets sujet-prédicat-objet ; en 2008, il est complété, entre autres, par le langage d'interrogation SPARQL (Sparql Protocol and RDF Query Language). Le standard OWL est présenté comme ayant l'avantage non seulement de permettre le traitement des informations en décrivant des classes et des propriétés, mais également de permettre la comparaison entre elles, c'est-à-dire d'autoriser l'exploitation de la



formulation exhaustive et rigoureuse d'un domaine de connaissance. Avec lui, c'est toute la question des ontologies qui est posée, une question qui apparaît comme centrale en ce qu'elle cristallise des oppositions fortes en matière de conceptions de la connaissance, du langage et de la communication (c'est particulièrement le cas au sein des chercheurs francophones).

Bien évidemment, beaucoup de travaux antérieurs à la mise en circulation du projet de web sémantique offrent des outils précieux pour sa mise en œuvre. Les ontologies se sont fortement développées à partir des années 90 dans le cadre de l'élaboration des systèmes à base de connaissances, qui succédaient eux-mêmes aux systèmes experts. Ce développement a mis en évidence qu'il s'agissait bien de procéder à une conceptualisation d'un domaine (par définition des concepts et de leurs relations) sous la forme d'une théorie logique et avec un vocabulaire spécifique, de manière qu'elle puisse donner lieu à des traitements automatiques (Charlet, Bachimont, Troncy, 2004). Comme le remarque Zacklad (2007), c'est parce qu'elles se situent dans le prolongement des principaux langages de représentation des connaissances issus des sciences cognitives et de l'intelligence artificielle que les ontologies « visent également à décrire des concepts qui sont appréhendés comme des représentations mentales plus ou moins universelles ou comme des catégories a priori largement partagées ». Cependant, il en résulte justement, selon lui, la nécessité de s'orienter plutôt vers la constitution d'un web « socio-sémantique », autrement dit de privilégier une approche non-logiciste capable, par exemple, de mettre en évidence différents points de vue sur les classifications retenues au sein d'une collaboration entre divers acteurs. Pour d'autres spécialistes, il ne s'agit pas là d'une remise en cause des bienfaits du web sémantique car il n'y a rien dans les standards recommandés par le W3C ni dans le fonctionnement de cet organisme qui empêche d'étendre les formalismes proposés pour qu'ils intègrent des dimensions sociales, sémiotiques ou pragmatiques (Gandon, 2006).

On observe, en fait, une polarisation de la réflexion liée étroitement aux appréciations portées sur les langues naturelles et les langues artificielles. Ainsi, l'attitude adoptée au sein du courant de l'IEML (Information Economy MetaLanguage) consiste à « dépasser » aussi bien celles-ci que celles-là : selon Lévy (2010), « natural languages, as well as notation systems invented before the 21st century, are not appropriate for the current and future scale of the social categorization process ». D'où la volonté de proposer l'édification d'un espace sémantique sur le web qui serait plus puissant que le web sémantique envisagé par le W3C, à partir de l'adoption d'un système symbolique correspondant à la construction d'un protocole de formalisation trans-linguistique et inter-réseaux des concepts. Le langage IEML incorpore bien des traits communs aux langues naturelles, mais sa syntaxe est entièrement régulière et son lexique exclut toute synonymie et toute homonymie, l'objectif étant de définir une topologie sémantique

sur des concepts formels. À l'inverse, il est possible de considérer que l'orientation souhaitable est celle qui consiste à s'intéresser à la sémantique du web plutôt qu'au web sémantique. C'est ce que préconise notamment Rastier (2008), qui introduit à trois types de débat avec les défenseurs des ontologies. Le premier de ces débats tient à ce qu'il n'est pas sûr que l'on puisse faire la liste des objets qui composent le monde ; le deuxième à ce que les connaissances sont sans doute plus inscrites dans des textes que dans des termes ; le troisième à ce qu'il est douteux qu'elles soient entièrement représentables par un formalisme logique. De ce point de vue, le web sémantique semble marqué par une prise en compte insuffisante des différences entre les langues, ainsi que de la diversité des discours et des genres de discours, ainsi que par le recours discutable à un petit nombre de relations sémantico-logiques assez pauvres. Dans le cas où l'on admet que les ontologies peuvent être utiles, il apparaît au moins que leur construction, lorsqu'elle s'effectue à partir de textes, implique la reconnaissance des difficultés de l'articulation entre le niveau linguistique et le niveau conceptuel, qui ne sont certainement pas le reflet l'un de l'autre (Mondary, Després, Nazarenko, Szulman, 2008). Cela a pour conséquence que l'établissement d'une première correspondance, entre des unités textuelles et des unités terminologiques, puis d'une seconde, entre ces unités terminologiques et les concepts, ne peut aucunement tabler sur une bijection entre les deux. D'une part, en effet, une seule unité terminologique peut renvoyer à plusieurs occurrences textuelles et des unités textuelles ambiguës sont susceptibles de renvoyer à plusieurs termes ; d'autre part, les termes qui vont être retenus pourront avoir des correspondant conceptuels en nombre variable, tandis que certains concepts n'auront pas de terme correspondant.

D'une certaine manière, la rencontre de telles situations ne fait que souligner l'importance du contexte. Les contenus des pages web sont vus, dans l'optique du web sémantique, comme des données dont un certain nombre de métadonnées seraient capables de représenter le sens afin qu'il soit manipulable par des traitements automatiques, alors que c'est précisément ce qui fait problème, comme l'a rappelé Beust (2009). On peut, en effet, penser que le sens d'un texte n'est pas exprimable en dehors d'un autre texte et qu'il ne peut donc relever que d'une activité d'interprétation. Depuis longtemps, c'est ce que rejettent, le plus souvent, les auteurs qui souhaitent l'élimination de toute ambiguïté dans la communication, réclament des discours qui disent la vérité du monde et vont, quelquefois, jusqu'à imaginer une langue qui échapperait à ce qui leur paraît un fléau. Mais est-ce un rêve ou un cauchemar ? En ayant envisagé l'un et l'autre, Orwell nous invite à faire preuve d'une certaine prudence en la matière.

### **Bibliographie**

BERNERS-LEE, T., HENDLER, J., LASSILA, O., 2001, « The semantic web », *Scientific American*, vol. 284, n° 5, pp. 35-43.

- BEUST P., 2009, « Ontologies et interprétation des textes sur le web », 16èmes Rencontres de Rochebrune : Ontologies et dynamiques des systèmes complexes, perspectives interdisciplinaires.
- CHARLET, J., BACHIMONT, B., TRONCY, R., 2004, « Ontologies pour le web sémantique », *Information - Interaction - Intelligence*, n° h. s. *Web sémantique*, pp. 69-99.
- CHURCHILL, W., 1943, discours à l'Université Harvard, 6 septembre. Repris dans CHURCHILL W., 1944, *Onwards to Victory : War Speeches*, Londres, Cassell.
- CHURCHILL, W., ROOSEVELT, F., 1944, lettres échangées entre avril et juin. Reprises dans KIMBALL W. F. (ed.), 1984, *Churchill and Roosevelt: The Complete Correspondence*, vol. 3, Princeton, Princeton University Press.
- CONDORCET, J. A. de, 1791, Second mémoire sur l'instruction publique, Paris, Imprimerie nationale.
- CONDORCET, J. A. de, 1792, Rapport et projet de décret sur l'organisation générale de l'instruction publique, Paris, Imprimerie nationale.
- COURTINE, J.-J., 1984, « George Orwell et la question de la langue », *L'Arc*, n° 94, pp. 54-60.
- CROISSET, A., 1903, « Les besoins de la démocratie », in LAVISSE É. et al., *L'éducation de la démocratie*, Paris, Alcan.
- GANDON, F., 2006, « Le web sémantique n'est pas antisocial », 17èmes Journées francophones d'ingénierie des connaissances, Nantes.
- LAKANAL, J., 1794, *Rapport sur l'établissement des écoles normales*, Paris, Imprimerie nationale.
- LANSON, G., 1903, « Etudes modernes dans l'enseignement secondaire », in LAVISSE É. et al., *L'éducation de la démocratie*, Paris, Alcan.
- LAUBLET, Ph., REYNAUD, Ch., CHARLET, J., 2002, « Sur quelques aspects du web sémantique », in LE MAITRE Cl. (ed.), *Actes des 2èmes assises nationales du GDR I3*, Toulouse, Cépaduès.
- LÉVY, P., 2010, « From social computing to reflexive collective intelligence: the IEMML research program », *Information Sciences*, vol. 180, p. 71-94.
- MONDARY, T., DESPRÉS, S., NAZARENKO, A., SZULMAN, S., 2008, « Construction d'ontologies à partir de textes : la phase de conceptualisation », 19èmes Journées francophones d'ingénierie des connaissances, Nancy.
- OGDEN, C. K., 1930, *Basic English: A General Introduction with Rules and Grammar*, Londres, Paul Treber.
- ORWELL, G., 1944, « As I please », *Tribune*, 18 août.
- ORWELL, G., 1946, « Politics and the English Language », *Horizon*, vol. 13, n° 76, pp. 252-265.
- ORWELL, G., 1949, *Nineteen Eighty-Four. A novel*, Londres, Secker & Warburg.

- PALMIER, J.-M., 1982, « Note sur l'apport du formalisme russe à l'analyse du langage politique », *L'homme et la société*, n° 63-64, pp. 95-108.
- RASTIER, F., 2008, « Sémantique du web vs semantic web ? Le problème de la pertinence », *Syntaxe & Sémantique*, n° 9, pp. 15-25.
- SÉRIOT, P., 1986, « La langue de bois et son double », *Langage et société*, n° 35, pp. 7-32.
- SLODZIAN, M., 2007, « Rationalisation des langues et terminologie : d'Ogden à Catford », *Hermès*, n° 49, pp. 61-68.
- TALLEYRAND, Ch. M. de, 1791, *Rapport sur l'instruction publique*, Paris, Imprimerie nationale.
- WORLD WIDE WEB CONSORTIUM (W3C) : [www.w3.org](http://www.w3.org).
- ZACKLAD, M., 2007, « Classification, thésaurus, ontologies, folksonomies : comparaisons du point de vue de la recherche ouverte d'information (ROI) », 35ème Congrès de l'Association canadienne des sciences de l'information, Montréal.