

Synonymie, hyponymie et thésaurus juridiques

par

Giancarlo TADDEI ELMI

Istituto per la Documentazione Giuridica del C.N.R. - ITALIA

845

DEFINITION DE THESAURUS

La nécessité de l'instrument thesaurus est née avec l'apparition des premiers systèmes de documentation automatique, et ce parce que, en raison de l'imprécision et de l'ambiguïté du langage naturel, tout système de documentation se heurte au problème suivant : l'ensemble des signes utilisés par la personne qui transmet un message correspond en partie seulement à l'ensemble des signes de la personne qui reçoit le message.

Il s'ensuit que, en réponse à une interrogation donnée, on a seulement une partie des documents utiles à la question et pouvant être recherchés ; d'autre part, seule une partie des documents pertinents est fournie à l'utilisateur, alors que bien souvent l'utilisateur se trouve avoir un "surplus" de documents ou d'informations non pertinents.

L'ensemble des documents pertinents, contenus dans le système, mais non recherchés est appelé "silence", tandis que l'ensemble des documents recherchés mais non pertinents est appelé "bruit" du système.

Tout système de documentation se pose le problème d'élaborer des stratégies de recherche visant à atténuer aussi bien le bruit que le silence, c'est-à-dire à améliorer la capacité de rappel et la capacité de précision des réponses.

Les thesaurus constituent un des instruments considérés comme les plus efficaces pour améliorer la qualité de la sélection des informations tant au niveau de rappel qu'au niveau de la précision.

Il existe un grand nombre de définitions de thesaurus, et on pourrait en donner bien d'autres, suivant tel ou tel aspect du thesaurus qu'on entend souligner. Ainsi, du point de vue de la fonction, on peut définir un thesaurus comme suit : "un instrument linguistique utilisé pour traduire en langage artificiel le langage naturel des documents et pour retourner au langage naturel à partir du langage artificiel". Du point de vue de la structure, on peut définir le thesaurus comme "un vocabulaire organisé et dynamique de termes ayant entre eux des relations de sens". (1)

Une des définitions qui, à mon avis, traduit le mieux la compréhension (l'aspect compréhensif) de ce concept est la suivante : "un ensemble ordonné de descripteurs constituant un système ouvert destiné à la classification et à l'organisation de concepts ou de termes par matière ou par problèmes ou bien selon d'autres critères". (2)

Cette définition a le mérite d'unifier les deux fonctions qui sont toujours présentes dans tout véritable thesaurus, à savoir la fonction de classification des concepts et la fonction d'organisation des concepts.

En tant que système de classification il tend à attribuer d'une façon univoque et réciproque, descripteur à concepts. En tant que système d'organisation, il tend au recensement des relations entre les concepts à travers une série de relations entre les descripteurs.

Cela signifie qu'un thesaurus se compose d'un ensemble de descripteurs ou de signes et d'un ensemble de relations entre ces descripteurs (signes).

Nous pouvons dire également que les relations entre les descripteurs sont définies de façon à obtenir une association univoque et réciproque entre le signe et son extension et sa compréhension.

Par extension, on entend l'objet, c'est-à-dire la classe d'éléments ou de choses réelles à laquelle le descripteur s'applique ou se réfère avec exactitude ; la compréhension indique au contraire le sens,

c'est-à-dire l'ensemble des attributs ou des propriétés qui caractérisent chaque élément auquel le descripteur s'applique correctement.

L'extension et la compréhension varient inversement l'une par rapport à l'autre : plus l'extension d'un terme descripteur est grande, plus sa compréhension est petite, et inversement. Par exemple, l'extension de *délit* est plus grande que celle de *vol simple*, car le premier terme se réfère à un plus grand nombre de choses ou entités; d'autre part, la compréhension de *vol simple* est plus grande que celle de *délit* car la caractérisation ou définition de *vol simple* doit faire état d'un plus grand nombre d'attributs que ceux qui suffisent à caractériser *délit*.

HOMONYMIE ET SYNONYMIE

Pour parvenir au premier résultat, c'est-à-dire à l'association univoque et réciproque entre concept et descripteur, il est nécessaire de résoudre les problèmes posés par les homonymies et les polysémies et de définir la synonymie; pour apporter une solution aux problèmes dérivant de l'homonymie, il est de règle de spécifier les descripteurs : par exemple dans le cas du mot *abattage*, qui peut avoir au moins trois significations - abattage des animaux, abattage d'arbres et abattage de mur - dans un thesaurus, il faudra créer un descripteur différent pour chaque signification.

La définition de synonymie est quelque peu complexe. Elle est un rapport paradigmatique de sens tendant à déterminer l'identité de sens entre deux unités de vocabulaire.

Une définition abstraite de synonymie pourrait être la suivante : deux termes sont synonymes lorsqu'ils ont la même compréhension, c'est-à-dire les mêmes composants de signification. Mais quand peut-on dire que deux termes ont le même sens en compréhension ?

Deux mots sont synonymes (ont la même compréhension) si les phrases résultant de la substitution de l'une par l'autre ont la même signification. Le point important de cette définition est que la synonymie est établie entre les mots et non entre les sens. La synonymie n'est qu'une partie du sens d'un mot (c'est le sens en compréhension, c'est-à-dire la signification).

En linguistique, cette notion de synonymie a été définie d'une manière rigoureuse sur la base de la notion d'implication réciproque de sens ou équivalence entre phrases (3).

Si une phrase S1 implique sémantiquement (implication directe) une autre phrase S2 et si S2 implique sémantiquement (implication converse) S1, alors S1 est sémantiquement équivalente à S2 ($S1 = S2$). Si S1 et S2 ont la même structure syntaxique et qu'elles se distinguent l'une de l'autre par le seul fait que, là où l'une contient un élément lexical x l'autre contient y, alors x et y sont synonymes seulement dans ces contextes (4).

En vue de l'établissement d'un thesaurus, il est de règle de distinguer deux types de synonymie, une synonymie au sens strict (appelée aussi parfaite, vraie ou totale) et une synonymie au sens large (appelée aussi imparfaite, partielle ou quasi-synonymie).

Nous avons la synonymie totale lorsque deux termes ou plus sont interchangeables dans tout contexte alors que nous avons la synonymie partielle lorsque deux termes sont interchangeables seulement dans des contextes déterminés ; on peut parler également de synonymie absolue opposée à synonymie relative.

SYNONYMIE ET LOGIQUE DES CLASSES

La synonymie peut être définie au moyen de relations entre classes et représentée à travers les diagrammes de Venn.

Soit x et y deux termes,

Soit A et B les classes ou extensions de x et de y , A et B peuvent être en relation d'équivalence ($A = B$), ou bien de non équivalence ($A \neq B$) :

a. équivalence (synonymie totale)

$$A = B$$

$$A = B$$

c'est-à-dire si et seulement si

$$A \subset B \text{ et } B \subset A$$

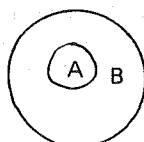
On a équivalence entre classes si et seulement si A est inclus dans B et B est inclus dans A .

On peut donc dire que, lorsque les classes de x et de y s'incluent, les termes x et y sont totalement synonymes ;

b. non équivalence (synonymie non totale et non synonymie)

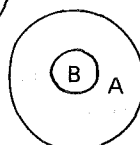
1^o cas : inclusion entre classes

$$A \subset B, \text{ et non } B \subset A$$



2^o cas : inclusion entre classes

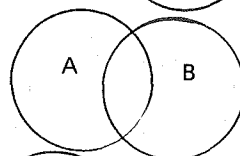
$$B \subset A, \text{ et non } A \subset B$$



3^o cas : intersection entre classes

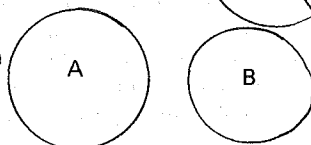
$$\text{non } B \subset A, \text{ non } A \subset B$$

$$A \cap B \neq \emptyset$$



4^o cas : non intersection

$$A \cap B = \emptyset$$



L'intersection entre les classes (cas n^o 3) représente la vraie synonymie partielle ; exemple : *demolizione* et *distruzione* (démolition et destruction), *pregiudizio* et *danno*.

Les cas 1 et 2 représentent en linguistique des cas d'hyponymie (ou d'inclusion) plus que de synonymie, même s'il ne me semble pas très juste de parler de synonymie partielle dans ces cas ; en effet, dans certains contextes, un terme inclusif peut très bien remplacer un terme inclus sans que change pour autant le sens du contexte. Par exemple : "Pierre a passé un accord avec Paul dans lequel Pierre s'engage à remettre un cheval à Paul et Paul à le payer". Dans ce contexte, *accord* est synonyme de *contrat de vente*. Il me semble qu'on

peut difficilement affirmer le contraire, c'est-à-dire que des termes inclus peuvent être synonymes de termes inclusifs.

Dans le cas n^o 4 on dit que l'intersection entre classes est vide et que, par conséquent, c'est un cas de non synonymie du point de vue linguistique.

SYNONYMIE DIRECTE ET CONVERSE

Le plupart des linguistes s'accordent à considérer que la synonymie absolue se rencontre rarement dans le langage naturel, qu'il soit courant ou technique.

La définition de la synonymie est en pratique variable et dépend de la matière qui est l'objet de la banque des données.

Ainsi, dans le droit français, *préjudice* et *dommage* sont des synonymes parfaits, et *licenciement*, *congédiement* et *congés* des synonymes imparfaits.

Dans le langage juridique italien, *pregiudizio* et *danno* sont des synonymes imparfaits et *licenziamento* et *congedo* ne sont jamais synonymes.

Dans la plupart des thesaurus, on distingue, à l'intérieur de la synonymie totale, deux types de synonymes, des synonymes préférentiels et de deuxième niveau. Les synonymes préférentiels sont reliés aux synonymes de deuxième niveau par deux rapports de renvoi, l'un direct et l'autre converse. Par exemple : *préjudice* employer *dommage* et *dommage* employé pour *préjudice*. Enfin on distingue la synonymie imparfaite de la synonymie parfaite en employant une relation d'association.

HYPONYMIE

La principale relation hiérarchique est la relation d'*hyponymie* ou d'*inclusion* (appelée également d'abstraction ou générique) entre concept ou terme supérieur (appelé *hyperonyme* ou superordonné) et concept ou terme inférieur (appelé *hyponyme* ou subordonné).

On l'appelle souvent inclusion. Il est toutefois préférable d'employer le terme hyponymie plutôt que le terme inclusion, car ce dernier est ambigu s'il se rapporte à des significations (l'inclusion entre significations est différente de l'inclusion entre classes et éléments d'une classe).

Par exemple, il est inexact de dire que la signification d'*automobile* est incluse dans celle de *véhicule*, ou que celle de *vente* est incluse dans celle de *contrat*. En effet, d'un certain point de vue un terme plus général est plus inclusif qu'un terme plus spécifique — *contrat* est plus inclusif que *vente*, *délit* est plus inclusif que *vol simple* — car il se rapporte à une classe plus étendue de choses ; mais d'un autre point de vue le terme plus spécifique est plus inclusif parce qu'il contient plus de composants de signification.

La différence de ces deux points de vue sur l'inclusion correspond dans la logique traditionnelle à la différence entre l'extension et la compréhension d'un terme.

C'est pourquoi, si l'on considère l'aspect en extension d'un terme, c'est-à-dire la classe d'éléments à laquelle il est applicable, dire que le générique inclut le spécifique est correct ; mais si l'on considère la compréhension, c'est-à-dire les attributs des éléments d'une classe, alors il est incorrect de dire que le spécifique est inclus dans le générique.

Cela dit, l'inclusion d'un terme plus spécifique dans un terme plus général peut s'exprimer également en termes de logique des classes et en termes d'ensembles : la classe ou l'ensemble des éléments pouvant être décrits comme *vente* ou auxquels on peut attribuer correctement le terme de *vente* est incluse

dans la classe ou l'ensemble des éléments pouvant être décrite comme *contrat*, et ainsi de suite.

Dans ce cas le terme d'inclusion est correct car il se rapporte à des éléments d'une classe et non aux significations. L'ensemble ou la classe des termes hyponymes constitue l'extension d'un concept (sens).

Une définition linguistique rigoureuse de l'hyponymie, utile également à des fins documentaires, est la définition en termes d'implication (sémantique) unilatérale (5). Le terme a (hyponyme) implique le terme b (hyperonyme) alors que le terme b n'implique pas le terme a ; si l'implication directe $a \rightarrow b$ est vraie, l'implication converse $b \rightarrow a$ n'est pas vraie et vice versa : prenons un exemple juridique : *vente* implique *contrat* tandis que *contrat* n'implique pas *vente*.

HYPONYMIE ET LOGIQUE DES CLASSES

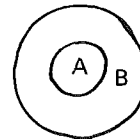
La relation d'hyponymie peut être définie au moyen de l'inclusion entre classes (extension des termes). L'inclusion entre classes peut être représentée, comme nous l'avons vu, par les diagrammes de Venn :

soit x et y deux termes,

soit A et B les classes ou extensions de x et de y,

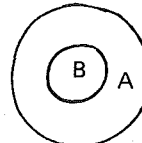
1^o cas : inclusion de A dans B (x est hyponyme de y)

$$\begin{aligned} A &\subset B \\ \text{non } B &\subset A \\ A &\neq B \end{aligned}$$



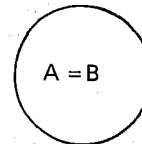
2^o cas : inclusion de B dans A (y est hyponyme de x)

$$\begin{aligned} B &\subset A \\ \text{non } A &\subset B \\ A &\neq B \end{aligned}$$



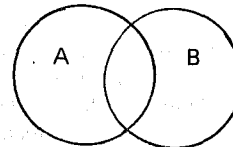
3^o cas : hyponymie bilatérale ou symétrique ou impropre (= synonymie totale)

$$A = B$$



4^o cas : non hyponymie (= synonymie partielle)

$$A \cap B \neq \emptyset$$



HYPONYMIE IMPROPRE ET SYNONYMIE

La notion d'hyponymie peut également servir à définir la synonymie. En effet, comme nous avons pu le voir à propos de la définition de synonymie en termes de classes, bien que généralement un terme hyperonyme n'implique pas (ou n'inclut pas) son hyponyme, il arrive souvent que le contexte le détermine dans le sens d'un de ses hyponymes ; en d'autres termes, il lui fait prendre la signification de l'hyponyme (cas 1^o et 2^o). Dans ce cas, nous sommes en général en présence de synonymies partielles dépendant des contextes. Quoi qu'il en soit, cela indique qu'il est possible de définir la synonymie totale comme hyponymie symétrique : si x est un hyponyme de y et si y est un hyponyme de x (c'est-à-dire si la relation est bilatérale ou symétrique) alors x et y sont synonymes (6). Cette hyponymie est appelée impropre sur la base de la théorie des

ensembles et de la logique des classes (cas 3^o).

La synonymie définie comme un cas d'hyponymie symétrique est une relation d'équivalence au sens mathématique du terme ; elle possède la propriété transitive caractéristique de l'hyponymie unilatérale ou asymétrique ; comme cas spécial d'hyponymie, elle possède la propriété symétrique et peut être définie également comme réflexive.

NOTES

- (1) ISO/DIS, Documentation-principes directeurs pour l'établissement et le développement de thesaurus monolingues destinés à la recherche documentaire, 1972.
- (2) Cfr. Leo Reisinger, Rechtsinformatik, W. de Gruyter, Berlin, 1977, p. 140.
- (3) Cfr. J. Lyons, Introduzione alla linguistica teorica, Bari, Laterza, 1971, pp. 588 et suivantes. Les rapports de sens peuvent être expliqués à partir des concepts d'assertion et de négation explicite. La phrase négative *le délit n'est pas grave* correspond à la phrase affirmative *le délit est grave* ; une phrase négative nie explicitement tout ce qui est affirmé explicitement par la phrase affirmative correspondante. Sur la base de cette notion d'assertion et de négation explicite, nous pouvons construire la notion sémantiquement plus utile et intéressante d'assertion et de négation implicite ou implication sémantique ou de sens.
- (4) Une autre manière de définir la notion d'équivalence sémantique pourrait être la suivante : si S1 et S2 impliquent sémantiquement chacune le même ensemble de phrases, alors elles sont équivalentes (cfr. Lyons, op. cit. p. 595).
- (5) Cfr. Lyons, op. cit. pp. 599 et suivantes.
- (6) Cfr. Lyons, op. cit. p. 602.