

La théorie glottogonique de Marr et l'isomorphisme structural entre les codes génétique et linguistique¹

Tamaz GAMKRELIDZE
Académie des sciences, Tbilissi

Mots-clés : ADN; code génétique; information; Jakobson; Marr; quatre éléments.

Dans mon exposé sur le problème de la théorie glottogonique de N. Marr, je veux aborder le problème des relations et des analogies entre les structures linguistiques et génétiques, en rapport direct avec le thème discuté.

Dans les années 1950, on a fait en biologie moléculaire une découverte de la plus haute importance, qui a fait la lumière sur le mécanisme de l'hérédité. On a découvert que l'hérédité correspond à une information inscrite tout au long des chromosomes, à l'aide d'un alphabet chimique défini.

En tant qu'éléments initiaux de cet alphabet, ou «lettres», on emploie quatre radicaux chimiques qui, en combinaison l'un avec l'autre dans la séquence linéaire des acides nucléiques, créent le texte chimique de l'information génétique. De même que la phrase forme un segment défini dans un texte linguistique constitué à l'aide de la séquence linéaire d'un certain nombre d'unités discrètes – des lettres ou des phonèmes, de même un gène particulier est conforme à un certain segment dans la longue chaîne des acides nucléiques, représentant les quatre radicaux chimiques. De même que, dans le code linguistique ces éléments initiaux — les phonèmes — sont privés de sens mais servent à former à l'aide de certaines combinaisons des unités minimales exprimant le contenu défini dans le système donné, d'une façon similaire, dans le code génétique ce n'est pas un élément isolé, un certain radical chimique, qui est informatif, mais les combinaisons particulières de ces quatre nucléotides initiaux par trois éléments, nommées «triplets», qui portent l'information.

¹ Ce texte est le résumé de l'exposé de T. Gamkrelidze, dont le texte définitif ne nous est pas parvenu. [Note de l'éditeur]

Il est possible de former seulement 64 combinaisons de quatre éléments initiaux par trois. C'est-à-dire que le «vocabulaire» génétique se compose de 64 «mots», dont les trois «triplets» se présentent comme les «signes de ponctuation», marquant dans la longue succession des acides nucléiques le début et la fin de la «phrase», et les autres correspondent à l'un des 20 aminoacides. En même temps, parmi les triplets de cette sorte on peut distinguer des «mots synonymiques», c'est-à-dire des suites qui correspondent aux mêmes aminoacides. L'établissement de telles corrélations entre les triplets de quatre éléments initiaux et de 20 aminoacides et la conversion de la longue chaîne des triplets en succession protéique des aminoacides, en enchaînement de peptide, est le déchiffrement et le décodage de l'information héréditaire contenue dans le code génétique, comme, par exemple, un message codé en alphabet morse est déchiffré par traduction en une autre langue.

Il devient évident que tout ce qui est vivant sur la terre possède la *connaissance* du code génétique au sens qu'il peut déchiffrer justement les *mots* constituant le contenu de l'information génétique et les synthétiser en succession protéique.

Ainsi, la variété infinie de tout le vivant peut se ramener aux très longs «messages» génétiques, formés des règles particulières de la combinaison linéaire des éléments du code génétique, possédant les traits frappants d'une ressemblance structurale avec le code linguistique. Ce n'est pas par hasard que dès le moment où l'on a déchiffré le code génétique, la génétique moléculaire vient emprunter abondamment les concepts et la terminologie linguistiques en décrivant le mécanisme de l'hérédité.

Pourtant, le nombre initial des unités linguistiques — phonèmes —, dont les combinaisons constituent les éléments significatifs minimaux d'une langue naturelle, est plus que quatre, et c'est le trait caractéristique du code linguistique, qui est à la base des langues et qui le distingue du code génétique. C'est ce qui crée une *redondance* dans le système linguistique, qui permet de corriger et de reconstruire les messages déformés par des facteurs extérieurs.

L'isomorphisme structural entre ces deux différents systèmes informatifs — génétique et linguistique — repose sur la combinaison linéaire des unités discrètes initiales, et cela pose une question phénoménologique sur la nature de ces systèmes et la cause de leur isomorphisme structural. Plusieurs points de vue sont avancés, que nous présenterons succinctement.

L'isomorphisme structural relevé entre les deux codes porte-t-il sur le caractère extérieur surgissant d'un rapprochement structural et de la convergence de deux systèmes divers qui réalisent les fonctions informationnelles analogiques, ou bien cet isomorphisme est le résultat du principe philogénétique de modelage du code linguistique à la manière et aux principes structuraux du code génétique? Cette dernière supposition est défendue par le linguiste Roman Jakobson, tandis que le généticien François Jacob pense que la structure analogique s'est formée par des systèmes différents de la fonction analogique.

La conception jakobsonienne de l'isomorphisme structural entre les codes génétique et linguistique suppose le processus évolutif de la superposition du code linguistique directement au code génétique et le copiage de ses principes structuraux, qui est réalisé dans les conditions où un être vivant possède inconsciemment la connaissance du caractère de ce dernier. Cela concerne entièrement la sphère de l'inconscient, en tant que l'organisme possède l'information sur la construction et la structure de ses mécanismes essentiels. Cela se relève non seulement dans le processus philogénétique de la formation des structures du mécanisme linguistique à la façon du code génétique, mais en même temps à l'action créatrice des personnalités éminentes, construisant les systèmes informationnels particuliers à la manière du code génétique sans la connaissance explicite de la structure de ce dernier.

Il convient d'évoquer, sous ce rapport, la théorie du processus glottogonique de l'éminent savant et linguiste N. Marr, qui était doué d'une fine intuition scientifique, le menant parfois jusqu'à la solution de certains phénomènes, bien que logiquement mal fondées, mais s'engageant en fin de compte dans la bonne direction.

N. Marr ramène, comme on sait, la diversité des langues historiques aux quatre éléments initiaux qui se composent, assez étrangement, de certains « triplets » de sons — des séquences de trois sons dénués de sens — *sal, ber, jon, roš*. Tous les textes de n'importe quelle longueur dans toutes les langues du monde sont le résultat de la transformation phonétique de ces quatre éléments initiaux dans la séquence linéaire définie, n'ayant eux-mêmes aucune signification. C'est de cette façon que, pour Marr, est déterminé l'unité du processus glottogonique.

La théorie glottogonique de Marr n'a pas une explication rationnelle, elle est en contradiction avec la logique de la linguistique théorique et sa méthodologie générale. En ce sens elle est irrationnelle et incongrue par rapport à la linguistique proprement dite. Mais cette théorie, présentant le modèle structural original de la langue qui est tout près du code génétique, n'est pas irrationnelle et incongrue par rapport à la psychologie et à la science en général, et elle peut illustrer la manifestation dans un savant d'une idée intuitive et inconsciente de la structure du code génétique. Marr ne pouvait pas évidemment posséder la connaissance explicite et consciente de la structure du système génétique, ainsi que ne la possédaient pas les philosophes Chinois qui ont élaboré, il y a trois mille ans, le système particulier des transformations de quatre éléments binaires composés du « principe masculin » *yang* et du « principe féminin » *yin*, se groupant en trois et donnant très exactement 64 triplets successifs, qui sont analogiques aux triplets génétiques. Dans le système symbolique des anciens Chinois, la diversité de tout ce qui est vivant et les rapports entre eux sont décrits à l'aide de la combinaison de ces mêmes triplets. De ce point de vue, les systèmes analogiques aux quatre éléments mondiaux dans la cosmologie des Ioniens, aux quatre états du corps humain chez Hippocrate semblent extrêmement significatifs, ainsi que les autres systèmes symboliques réali-

sant l'idée qu'une relation stricte est imposée par une filiation inconsciente entre des systèmes sémiotiques de telle sorte et le code génétique. Tous ces systèmes symboliques, comme le modèle de la langue de Marr, coïncident d'une manière frappante, jusqu'aux paramètres quantitatifs, avec la structure du code génétique jouant dans leurs créateurs évidemment le rôle d'un substrat modelant inconscient.