

De la composition comme dispositif analogique

par Philippe Monneret*, Mariangela Albano**

Introduction

La nature plurivoque du signe linguistique donne lieu à plusieurs parcours possibles dans la construction des signifiés. Cette plasticité du langage humain permet de répondre de différentes façons aux besoins lexicaux qui ne peuvent être satisfaits par la récupération en mémoire d'unités lexicales mémorisées et par conséquent requièrent une création néologique. Outre l'emprunt, les deux solutions majeures impliquant l'organisation interne de la langue sont la polysémie et la composition. Ces deux procédés se caractérisent notamment par leur caractère économique au plan cognitif: la polysémie utilise un mot disponible en le pourvoyant d'une nouvelle acception (ex. *souris* "rongeur" vs *souris* d'ordinateur) tandis que la composition, ou plus largement la phraséologie, joue sur une combinatoire particulière, parce qu'elle n'est pas strictement compositionnelle, d'unités lexicales disponibles (ex. *voiture-bélier*).

1. La composition comme procédé de formation de mots

Classiquement, compte non tenu des autres procédés de formation lexicale qui ne sont pas pertinents pour notre propos, la composition se distingue de la dérivation par le caractère autonome de ses éléments constitutants – qui s'oppose à l'absence d'autonomie des préfixes et des suffixes¹: «un mot-forme est formé par composition lorsqu'il résulte

* Université Paris Sorbonne, Paris IV.

** Université Sorbonne Nouvelle, Paris III.

¹ Bien entendu, certains cas, qui restent marginaux, prêtent à discussion. Par

de la concaténation – c’est-à-dire de la juxtaposition linéaire – de plusieurs mots-formes ou radicaux» (Polguère, 2008, p. 79).

Au plan graphique, les éléments constitutants du mot composé français peuvent être soudés. Dans ce cas, la soudure est ou bien complète et le composé s’écrit en un seul mot (ex. *portefeuille*) ou bien marquée par un trait d’union: *rouge-gorge* (adjectif + nom), *sourd-muet* (adjectif + adjectif), *timbre-poste* (nom + nom), *porte-plume*, *réveille-matin* (verbe + nom), *après-midi*, *sous-sol* (préposition + nom). Mais, dans de nombreux cas, le nom composé ne se signale graphiquement par aucune caractéristique particulière: *boîte à chapeaux*, *eau de vie* (nom + préposition + nom), *assurance vieillesse* (nom + nom).

La question qui se pose alors est celle de la différence entre une “concaténation” formant un mot composé et une construction syntaxique bien formée quelconque. Or il semble exister aujourd’hui un consensus assez large en faveur d’une conception continuiste de la composition: c’est le degré de figement des éléments d’une séquence qui permet de distinguer le mot composé d’une séquence libre d’unités². Il n’y a donc pas de différence structurelle entre un “mot composé” et une séquence plus ou moins figée; autrement dit, le concept de composition se résorbe intégralement dans la problématique de la phraséologie, dont le figement est un des critères centraux³.

exemple, les éléments de type *super* (ex. *supermarché*) peuvent être considérés comme des préfixes mais, en raison de la possibilité qu’ils ont d’être employés d’une manière autonome («c’est super!»), ils peuvent aussi être considérés comme des éléments de noms composés. Pour des raisons que nous ne pouvons pas développer ici, la première solution nous semble préférable.

² On notera cependant que certains auteurs souhaitent maintenir une différence entre mot composé et séquence figée, tout en reconnaissant que la frontière entre les deux catégories n’est pas très nette. Polguère, par exemple, considère qu’«il existe un certain flottement dans la façon dont s’orthographient les lexies composées (souvent appelées mots composés); ce flottement est le reflet du caractère flou de la frontière qui sépare parfois en français locutions et lexies composées» (Polguère, 2008, p. 80). Il fait aussi remarquer que dans la majorité des manuels la composition est traitée à côté de la lexicalisation des syntagmes comme, par exemple, *fait divers*, *pomme de terre* ou *laisser tomber*.

³ Le terme utilisé par Benveniste pour décrire ce type de phénomènes est celui de *synapse*: «un “aigle pêcheur à tête blanche” pourrait être un syntagme descriptif de création occasionnelle. Mais s’il est donné dans la nomenclature zoologique comme le nom, usuel dans une certaine communauté, d’un oiseau particulier, alors *aigle pêcheur à tête blanche* devient en français de Nouvelle-Calédonie une dénomination synaptique à enregistrer dans le lexique de cette science et de cette communauté» (Benveniste, 1966, t. 2, p. 175).

Une autre distinction usuelle dans l'analyse de la composition en français doit être mentionnée: la différence entre les mots composés d'origine populaire et les mots composés d'origine savante. Dans le premier cas, il s'agit d'un mot composé qui montre «des bases françaises et les assemble suivant l'ordre de la syntaxe française» (Mortureux, 2013 [2004], p. 53). Quant aux composés savants, ils sont formés de mots empruntés aux langues classiques, notamment au grec et au latin.

Mortureux, dans *La lexicologie entre langue et discours*, consacre un chapitre entier aux composés populaires et savants. Elle distingue, parmi les composés populaires, les “composés à base verbale” et les “composés à base nominale”. Le premier type résulte d'une nominalisation d'une séquence «conservant l'ordre des mots français, qui place le verbe devant son complément» (ivi, p. 54). Ainsi, le composé *compte-gouttes* est formé sur deux bases, l'une verbale, l'autre nominale, et présente l'ordre de la phrase française. Même sur le plan sémantique, le composé est caractérisé par un caractère “partiellement compositionnel” puisque son sens est prédictible et “partiellement conventionnel” dans la mesure où la séquence est conçue comme une unité. Quant aux composés à base nominale, Mortureux considère qu'il convient de distinguer deux types de composés: les composés endocentriques et les composés exocentriques⁴:

les composés endocentriques (tels que *porte-fenêtre* ou *procès-verbal*) sont de même niveau syntaxique que leur base sujet et attribut du sujet; à l'inverse, les composés exocentriques (*peau-rouge*, *pur-sang*) sont le produit de la nominalisation d'un constituant de niveau inférieur (complément d'objet du verbe «avoir» et attribut de ce complément de ce complément dans les deux exemples ci-dessus) (ivi, p. 55)⁵.

⁴ Ces termes sont empruntés à la syntaxe et ont été introduits par Bloomfield (1970 [1933]) pour distinguer les deux patrons de formation que suivent les constructions syntaxiques. D'une part, l'endocentrisme est le mécanisme permettant de construire un syntagme composé de constituants immédiats dont la résultante appartient à la même classe formelle que l'un d'entre eux. D'autre part, une séquence endocentrique est formée de composants liés entre eux par un rapport de subordination ou de coordination, sans changement de catégorie lexicale ou grammaticale. L'exocentrisme, en revanche, met en présence des éléments appartenant à des catégories différentes pour en former une nouvelle. Une construction exocentrique est représentée par une suite d'éléments formant entre eux une nouvelle classe formelle, différente de celle des constituants. C'est d'ailleurs ce critère qui est utilisé par certains auteurs (notamment Polguère 2008) pour isoler la composition parmi l'ensemble des unités phraséologiques: le terme “composition” est dans ce cas réservé aux unités phraséologiques endocentriques. Ce critère, on l'aura compris, nous semble insuffisamment justifié pour être définitivement retenu.

⁵ Les italiques sont dans le texte.

Il existe donc des opérations lexico-syntaxiques complexes dans le procédé de composition des mots. Selon cette perspective, dans les termes de Benveniste (1966, t. 2, p. 161), il existe un rapport entre la composition et la syntaxe. Benveniste, lui-même, après avoir analysé la composition nominale, qualifie ce procédé comme le produit «des constructions syntaxiques avec leur variété de prédication» (*ibid.*). L'auteur admet que la fonction d'un mot composé est celle de «transférer au virtuel le rapport actuel de prédication énoncé par la proposition de fondement» (*ibid.*).

Mortureux, dans son ouvrage qui prolonge sur plus d'un point les analyses de Benveniste, reconnaît que la composition se situe entre syntaxe et lexique.

D'une part, le contenu de cette prédication s'inscrit dans la syntaxe; d'autre part, la composition implique également une dimension "collective" (ou normative) qui caractérise le lexique: «le nom composé, qui, comme tout lexème, est un signe, associe [un] concept (qui devient le signifié du signe) à une forme (qui devient son signifiant), et, en l'inscrivant dans le lexique de la langue, il lui assure un reconnaissance collective» (Mortureux, 2013 [2004], p. 55).

Quant à la différence entre les composés populaires et les composés savants, elle est relative à la nature des unités composantes ainsi qu'à certaines règles de composition: «les composés savants présentent [...] une 'voyelle de liaison' à la jonction des bases, -o ou -i, selon le modèle greco-latin (*cancérogène*⁶, *anxiogène* dans le vocabulaire médical et *insecticide*, *pesticide* dans celui de la chimie), et l'ordre interne classique déterminant-déterminé» (ivi, p. 56)⁷.

Les éléments formatifs des composés savants sont des bases non autonomes comme -gène, -cide, -logie, -logie qui sont intégrées à l'ensemble des morphèmes lexicaux du français et recouvrent une large partie du lexique spécialisé.

Mortureux précise qu'il existe «des mots composés associant une base d'origine grecque et une base d'origine latine (*hydravion*), ou encore une base française et une base d'origine grecque (*monoplace*) ou latine (*biplace*)» (*ibid.*)⁸. Ces bases non autonomes sont souvent assimilées à des affixes et on prend rarement en compte leur dimension sémantique qui, au contraire, nous révèle leur valeur dénomminative.

Ajoutons enfin, pour achever ce rapide tour d'horizon, que, outre la dérivation et la composition, qui demeurent les procédés d'enrichis-

⁶ On notera que cette forme est en concurrence avec la forme *cancérigène*.

⁷ Les italiques et les guillemets simples sont dans le texte.

⁸ Les italiques sont dans le texte.

sement lexical les plus étudiés, d'autres phénomènes en rapport avec la composition doivent être pris en compte: l'abréviation, la troncation, le mot-valise, la co-prédication, la siglaison et l'acronymie⁹.

2. La composition: entre Saussure et Bally

La question de la production d'unités linguistiques nouvelles renvoie à une interrogation sur la systématité de l'analogie. C'est ce sur quoi

⁹ Comme le remarque avec justesse Mortureux (2013 [2004], pp. 58-9), il reste possible de réactiver des mots composés de nature savante. En premier lieu, ils peuvent être abrégés comme dans le cas de *télévision* qui devient *télé*. En deuxième lieu, les bases issues par abréviation des noms composés représentent les bases pour former de nouveaux mots comme pour *télespectateur*. Dans cette perspective, deux phénomènes peuvent être relevés: la troncation et l'homonymie. On trouve parfois «une forme tronquée d'une ou des deux bases, sans que celle-ci se soit autant lexicalisée sous cette forme abrégée» (*ibid.*). Un exemple est *héliport*, forme tronquée de deux mots: *hélicoptère* et *aéroport*. Quant à l'homonymie, elle est impliquée dans la mesure où l'évolution des composés savants peut conduire à l'émergence d'«homonymes appartenant à des catégories lexicales différentes, bases autonomes et éléments de mots construits» (*ibid.*). C'est par exemple le cas d'*auto-* qui peut représenter l'origine du pronom grec «soi-même» dans *autonome* ou *autodidacte* mais il peut aussi être la réactivation de la forme abrégée *auto* (issue de *automobile*) dans *auto-école* ou *autodidio*. Enfin, cette réversibilité des bases savantes est caractérisée par la «démotivation», qui est un processus produisant des unités composées de plusieurs abréviations ou troncations qui ont perdu leur motivation. Ainsi, Mortureux (*ibid.*) donne l'exemple de l'adjectif français *télégénique*, qui est composé de *télé-* provenant du nom *télévision* et du suffixe *-génique* dont l'origine est *photogénique*. Par ailleurs, l'évolution de la langue française a permis l'émergence de nouvelles formes de composition servant à limiter la longueur des mots. Dans cette catégorie, «mot-valise», est une forme de «télescopage de deux bases, dont chacune est tronquée, mais dans des conditions telles que le mot crée conserve un segment commun aux deux bases» (ivi, p. 60). Ainsi, le mot *informatique* est le résultat de la rencontre du nom *information* et de l'adjectif *automatique*. Un phénomène similaire est représenté par des mots comme *photocopillage* ou *télématique*. Dans le premier cas, il s'agit d'une forme de troncation savante qui réactive un mot-valise déjà existant (*photocopie*) avec un autre substantif, *pillage*. En revanche, le deuxième cas montre une modalité de formation «qui tient à la fois, pour les signifiants, de la composition savante par troncation, et pour les signifiés, de la co-prédication» (ivi, p. 62). Ainsi, *télématique* regroupe le mot *télécommunication* et le mot-valise *informatique*. Les derniers cas qui se rattachent à la composition sont la siglaison et l'acronymie. Dans le premier cas, on prend en compte l'abréviation d'une suite des mots formée par les initiales comme, par exemple, CAPES pour «Certificat d'Aptitude Professionnelle à l'Enseignement Secondaire». Dans le second, les acronymes offrent la possibilité de travailler sur le résultat du sigle en évitant les homonymies ou un problème de prononciation. C'est le cas de «COURLY» (ivi, p. 63) explicitant la séquence «COmmunautéURbaine de LYon».

insiste à juste titre Saussure (1967 [1916], pp. 221-37 et 242-5), qui décrit l'analogie comme un procédé qui «s'exerce en faveur de la régularité et tend à unifier les procédés de formation et de flexion» (ivi, p. 222). Une forme dite "analogique" est:

une forme faite à l'image d'une ou plusieurs autres d'après une règle déterminée. Ainsi le nominatif latin *honor* est analogique. On a dit d'abord *honōs*: *honōsem*, puis par rotacisation de l's *honōs*: *honōrem*. Le radical avait dès lors une double forme; cette dualité a été éliminée par la forme nouvelle *honor*, créée sur le modèle de *ōrātor*: *ōrātorem*, etc., par un procédé [...] que nous ramenons dès maintenant au calcul de la quatrième proportionnelle *ōrātorem*: *ōrātor* = *honōrem*: *x*, *x* = *honor*[...] pour contrebalancer l'action diversifiante du changement phonétique (*hōnos*: *honōrem*), l'analogie a de nouveau unifié les formes et rétabli la régularité (*honor*: *honōrem*)¹⁰ (ivi, pp. 221-2)¹¹.

L'analogie apparaît ainsi comme un processus de régularisation visant à compenser les désorganisations, dans les systèmes de la langue, produites par les évolutions phonétiques. Dans le cas de l'évolution de *honos*, cette désorganisation se manifeste par l'absence de parallélisme du couple *honos* / *honorem* (par opposition au parallélisme illustré par *orator* / *oratore*), provoqué par la rotacisation. L'analogie est conçue comme un phénomène psychologique et grammatical car «elle suppose la conscience et la compréhension d'un rapport unissant les formes entre elles» (ivi, p. 226).

Par ailleurs, pour rendre compte de la formation de mots à partir de plusieurs unités lexicales, Saussure utilise le concept d'agglutination, tendance des mots à se souder en unités inanalysables:

l'agglutination consiste en ce que deux ou plusieurs termes originellement distincts, mais qui se rencontraient fréquemment en syntagme au sein de la phrase, se soudent en une unité absolue ou difficilement analysable. Tel est le processus agglutinatif: processus, disons-nous, et non procédé¹², car ce dernier mot implique une volonté, une intention, et l'absence de volonté est justement un caractère essentiel de l'agglutination (ivi, p. 242).

¹⁰ Les italiques sont dans le texte.

¹¹ Saussure résume le fait analogique dans une formule efficace. Il affirme que «tout fait analogique est un drame à trois personnages, qui sont: 1° le type transmis, légitime, héréditaire (par exemple *honōs*); 2° le concurrent (*honor*); 3° un personnage collectif, constitué par les formes qui ont créé ce concurrent (*honōrem*, *ōrātor*, *ōrātōrem*, etc.)» (Saussure, 1967 [1916], p. 224).

¹² La distinction entre processus et procédé est intéressante. Elle montre bien la dimension non-consciente du phénomène. Comme on le verra plus loin, les deux processus fondamentaux de l'analogie et de la fusion sont le plus souvent non-conscients.

L'agglutination peut souder deux mots comme *tous* et *jours* qui deviennent *toujours* ou *au jour d'hui* qui se transforme en *aujourd'hui* ou encore des sous-unités d'un mot comme dans le cas du superlatif indo-européen **swād-is-to-s* cité par l'auteur (ivi, p. 243)¹³.

On notera au passage que Saussure oppose nettement le processus de l'agglutination au précédé de l'analogie:

1° Dans l'agglutination deux ou plusieurs unités se confondent en une seule par synthèse (par exemple encore, de hanchoram), ou bien deux sous-unités n'en forment plus qu'une (cf. hēd-isto-s, de **swād-is-to-s*). Au contraire l'analogie part d'unités inférieures pour en faire une unité supérieure. Pour créer pāg-ānus, elle a uni un radical pāg- et un suffixe -ānus.

2° L'agglutination opère uniquement dans la sphère syntagmatique; son action porte sur un groupe donné; elle ne considère pas autre chose. Au contraire l'analogie fait appel aux séries associatives aussi bien qu'aux syntagmes.

3° L'agglutination n'offre surtout rien de volontaire, rien d'actif; [...] c'est un simple processus mécanique, où l'assemblage se fait tout seul. Au contraire, l'analogie est un procédé, qui suppose des analyses et des combinaisons, une activité intelligente, une intention (ivi, p. 287).

Saussure indique également que les constructions par agglutination montrent «la cimentation lente d'éléments qui, en contact dans un syntagme, ont subi une synthèse pouvant aller jusqu'au complet effacement de leurs unités originelles» (*ibid.*), tandis que les constructions de nature analogique représentent un «agencement obtenu d'un seul coup, dans un acte de parole, par la réunion d'un certain nombre d'éléments empruntés à diverses séries associatives» (*ibid.*).

Mais cela ne signifie aucunement que la formation des mots composés dépendrait du processus d'agglutination et que, par conséquent, l'analogie n'aurait aucun rapport avec la composition. Car Saussure différencie le mot *composé* du mot *agglutiné*, et si l'agglutination con-

¹³ Selon Saussure, l'agglutination comprend trois phrases distinctes. La première représente la «combinaison de plusieurs termes en un syntagme, comparable à tous les autres» (Saussure, 1967 [1916], p. 243); la deuxième phrase est appelée «agglutination» car elle est la synthèse des éléments du syntagme en une unité nouvelle. Saussure affirme que «cette synthèse se fait d'elle-même, en vertu d'une tendance mécanique: quand un concept composé est exprimé par une suite d'unités significatives très usuelle, l'esprit, prenant pour ainsi dire le chemin de traverse, renonce à l'analyse et applique le concept en bloc sur le groupe de signes qui devient alors une unité simple» (*ibid.*). Enfin, la troisième est caractérisée par «tous les autres changements susceptibles d'assimiler toujours plus l'ancien groupe à un mot simple: unification de l'accent (vért-jús → verjús), changements phonétiques spéciaux, etc.» (*ibid.*).

cerne exclusivement le second, l'analogie est bien le mécanisme responsable de la formation des mots composés:

On voit combien il importe de distinguer l'un et l'autre mode de formation. Ainsi en latin *possum* n'est pas autre chose que la soudure de deux mots *potissum* «je suis le maître»: c'est un agglutiné; au contraire, *signifier*, *agricola*, etc., sont des produits de l'analogie, des constructions faites sur des modèles fournis par la langue. C'est aux créations analogiques seules qu'il faut réserver les termes de composés et de dérivés. [En notes] Ceci revient à dire que ces deux phénomènes combinent leur action dans l'histoire de la langue; mais l'agglutination précède toujours, et c'est elle qui fournit des modèles à l'analogie (ivi, p. 244).

Les mots composés et dérivés sont donc pour Saussure considérés comme des produits d'un procédé¹⁴ analogique. Saussure précise cependant que l'agglutination et l'analogie se combinent au niveau diachronique et que, par exemple, le composé grec *hippó-dromo-s* se forme par agglutination à partir des désinences de l'indo-européen «mais c'est l'analogie qui en a fait une formation productive avant la soudure absolue des éléments» (ivi, pp. 244-5): «c'est par l'intervention de l'analogie que l'agglutination crée des types syntaxiques et travaille pour la grammaire; livrée à elle-même, elle pousse la synthèse des éléments jusqu'à l'unité absolue et ne produit que des mots indécomposables et improductifs (type *hanchōram* → *encore*), c'est-à-dire qu'elle travaille pour le lexique» (*ibid.*).

Cette forme de régularité, au sens paradigmatique, dans les structures de la langue française, est également abordée par Bally (1951 [1909]) à partir d'une distinction entre l'instinct étymologique et l'instinct analogique, distinction qu'il applique à l'analyse des locutions composées.

L'instinct étymologique représente une force tendant «à relier entre eux les mots ou les éléments constitutifs des mots, en considérant leur forme extérieure et non leur rapport avec l'idée dont ils sont les symboles» (ivi, p. 32).

D'autre part, l'action analogique est une forme d'exagération de l'instinct étymologique (ivi, pp. 38-9): «l'instinct étymologique crée des sens inexistantes [...] l'instinct analogique crée des formes; au fond le résultat psychologique est le même» (*ibid.*). Par exemple, un étudiant étranger confronté à des formes verbales comportant le préfixe *re-* (qui indique généralement la répétition d'une action annoncée par le verbe) sera peut-être conduit à la création de sens inexistantes: «c'est son instinct étymologique qui l'induit en erreur: ainsi 'relire' est con-

¹⁴ Nous verrons plus loin qu'il s'agit plutôt d'un processus que d'un procédé.

forme au sens vivant du préfixe, mais ‘remarquer’ dans son sens le plus courant, ne se décompose pas en ‘marquer de nouveau’, ou plutôt il ne se décompose pas du tout» (ivi, p. 39).

Inversement, Bally indique que si cet étudiant est également susceptible de créer des mots inexistantes s’il généralise l’usage du préfixe *re-* et produit des formes non attestées. Dans ce cas, il «sera victime de l’analogie: ainsi *revenir*, *retourner*, *remonter*, etc. peuvent faire croire qu’on dit aussi *raller*, *refuir*, ce qui n’est pas»¹⁵ (*ibid.*)¹⁶.

L’analyse de Bally correspond assez bien à ce qu’en termes actuels on appellerait “pensée analogique” (Gentner, 1983; Clement, Gentner, 1991; Gentner, Holyoak, Kokinov, 2001; Hofstadter, 2001; Hofstadter, the Fluid Analogies Research Group, 1995; Hofstadter, Sander, 2013). En effet, ces deux “instincts” ne sont autres que des processus analogiques qui répondent au besoin – dicté par notre système cognitif – de créer ou d’interpréter des formes à partir de modèles existants, c’est-à-dire disponibles en mémoire. L’analogie, par sa fonction régularisatrice, établit des similarités structurelles.

Nous constatons donc que, chez Saussure comme chez Bally, l’analogie est sollicitée – plus ou moins explicitement – pour l’analyse des mots composés. Il nous reste maintenant à examiner de plus près cette notion d’analogie et à en faire apparaître la dimension cognitive.

3. Analogie et composition

De nombreuses recherches en psychologie (p. ex. Piaget, 1975) et en intelligence artificielle (p. ex. Evans, 2000), ont permis d’établir que l’analogie est un élément fondamental du système cognitif humain, en ce qu’elle permet d’identifier des structures relationnelles similaires dans des contextes différents pour transférer l’information d’un contexte à l’autre (Holyoak, Thagard, 1989, 1995), d’une source mémorisée ou connue, à une cible non encore connue ou non encore identifiée.

¹⁵ Les italiques sont dans le texte.

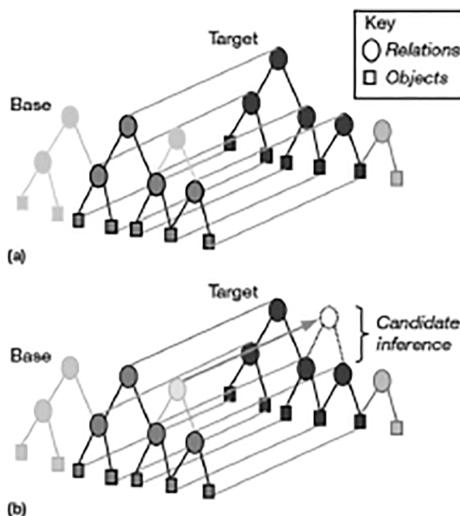
¹⁶ On observe cependant que ces deux verbes sont attestés par le dictionnaire du CNRTL: *refuir* avec le sens «fuir, s’enfuir» (<http://www.cnrtl.fr/definition/refuir>), *raller* avec le sens de «aller de nouveau (qq. part), repartir, retourner («qq. part» (<http://www.cnrtl.fr/definition/dmf/raller>)). Bally ne disposait certes pas des ressources électroniques qui sont aujourd’hui à notre disposition. On retiendra de cet exemple malencontreux qu’il convient évidemment d’effectuer les vérifications nécessaires avant d’affirmer l’inexistence d’une forme. Il reste cependant que, s’ils sont attestés, *raller* et *refuir* demeurent des verbes qui ne sont pratiquement plus en usage.

Depuis les travaux de Gentner (1983; Clement, Gentner, 1991; Markman, Gentner, 1993; Gentner, Holyoak, Kokinov, 2001; Gentner, Smith, 2012) qui font autorité dans ce domaine, la modélisation courante des processus analogiques est celle de la “structure mapping theory” (FIG. 1).

Elle se caractérise par le fait qu’une structure source (comportant des objets, des relations, et des relations entre des relations) est mise en correspondance structurelle (c’est le “mapping”, difficile à traduire en français) avec une structure cible. Cette mise en correspondance s’effectue d’abord sur les éléments les plus accessibles au parallélisme, mais peut être étendue à des parallélismes non accessibles de prime abord, qui sont inférés à partir de la structure source (voir FIG. 1, “candidate inference”).

FIGURE 1

La théorie du “Structure mapping” (Gentner, Smith, 2012)



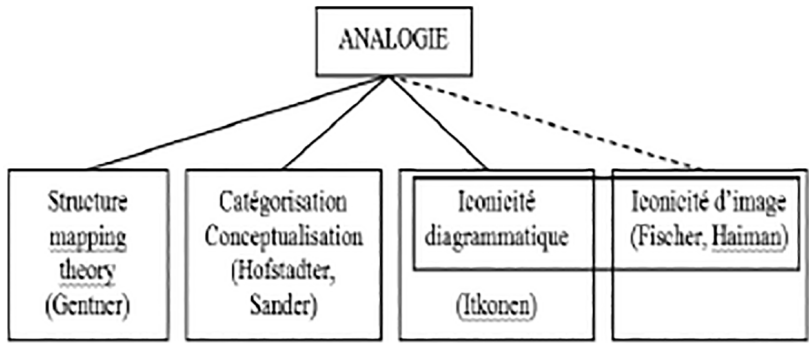
En psychologie, l’analogie a longtemps été considérée comme un simple cas particulier de raisonnement inductif, par opposition au raisonnement déductif. Ce processus était alors conçu seulement dans sa dimension consciente et ce n’est que depuis une vingtaine d’années que son horizon s’est élargi à la sphère non-consciente.

Aujourd’hui, l’analogie est considérée comme le processus central de la cognition humaine (“the core of cognition” selon l’expression de Douglas Hofstadter, 2001), en tant qu’il constitue le processus cognitif qui permet la catégorisation.

Les travaux actuels, dans le champ de la psychologie cognitive, visent à documenter différents aspects de l’analogie sous l’angle des fonctions exécutives (mémoire, inhibition, flexibilité cognitive), depuis leur développement chez l’enfant jusqu’à leur dégénérescence chez les personnes âgées.

En linguistique, les travaux de Monneret (2004, 2014, 2017), à la suite d’Itkonen (2005), ont cherché à documenter les contreparties langagières des processus analogiques. L’ensemble des champs couverts par la problématique analogique en linguistique (ou “linguistique analogique”) est représenté par la figure 2 ci-dessous:

FIGURE 2
Les problématiques analogiques en linguistique



Outre les questions relatives à la catégorisation et à l’iconicité diagrammatique dont les rapports avec l’analogie ont déjà été étudiés par différents auteurs, l’iconicité d’image¹⁷ – dont un cas remarquable

¹⁷ Le concept d’iconicité est emprunté à Peirce, via Jakobson (1965). Il peut être défini comme une analogie hétérogène (par opposition à la catégorisation qui constitue un cas d’analogie homogène) entre forme et contenu, ou entre signifié et signifiant. Olga Fischer définit l’iconicité de la façon suivante: «Iconicity as a semiotic notion refers to a natural resemblance or analogy between the form of a sign (‘the signifier’, be it a letter or sound, a word, a structure of words, or even the absence of a sign) and the object or concept (‘the signified’) it refers to in the world or rather in our perception of the world» (Fischer, 2000). L’iconicité d’image se distingue de l’iconicité diagrammatique comme l’analogie binaire se distingue de l’analogie proportionnelle: «The similarity between sign and object may be due to common features inherent in both: by direct inspection of the iconic sign we may glean true information about its object. In this case we speak of ‘imagic’ iconicity (as in a portrait or in onomatopoeia,

est celui du symbolisme phonétique – peut être également considérée comme faisant partie des phénomènes linguistiques conditionnés par l’analogie (Monneret, 2014).

Dans cette perspective, l’analogie se définit comme un processus cognitif d’identification qui repose sur une relation de similarité (binaire ou proportionnelle) entre deux entités (ou plus), relation établie, consciemment ou non, par un individu singulier dans un contexte singulier. Bien entendu, comme l’indique le tableau 1 ci-dessous, il convient également de prendre en compte le plan collectif des langues, qui résulte d’effets de sommation à partir des processus individuels.

TABLEAU 1
L’analogie au plan individuel et au plan collectif

	Processus analogiques	Structures analogiques
Niveau individuel	Processus analogique réalisé par le système cognitif du locuteur réel	Structure analogique disponible dans le système cognitif du locuteur réel, fondée sur des régularités ou similarités phonologiques, morphologiques, syntaxiques, lexicales, pragmatiques, perceptibles ou actualisées inconsciemment par le locuteur réel (donc stockées dans sa mémoire)
Niveau collectif	Processus analogique validé au plan d’une langue (analogie dans le changement linguistique, compte tenu des normes présentes dans chaque synchronie considérée)	Structure analogique disponible dans une langue* donnée (qui résulte d’un phénomène de cognition sociale impliquant des normes), matérialisée par des régularités ou similarités phonologiques, morphologiques, syntaxiques, lexical, pragmatique, etc. observables par le linguiste

* Ou un dialecte, un sociolecte, ou toute autre stabilisation collective d’un système linguistique.

e.g. ‘cuckoo’) and the sign is called an ‘iconic image.’ [...] When we have a plurality of signs, the analogy may be more abstract: we then have to do with diagrammatic iconicity which is based on a relationship between signs that mirrors a similar relation between objects or actions (e.g. a temporal sequence of actions is reflected in the sequence of the three verbs in Caesar’s dictum “veni, vidi, vici”): in this instance, the sign (here the syntactic structure of three verbs) is an ‘iconic diagram’».

Comment la question de la composition peut-elle être prise en charge dans une perspective analogique? Il convient tout d'abord de distinguer plusieurs *fonctions* de l'analogie, dont certaines découleront naturellement des remarques précédentes.

Les principales fonctions de l'analogie sont les suivantes: la fonction *catégorisatrice* (sollicitée pour tout processus de catégorisation), la fonction *régularisatrice* (qui vise à accroître la systématité des langues en régularisant les paradigmes), la fonction *iconique* (grâce à laquelle un signe gagne en efficience cognitive¹⁸), la fonction figurative (à l'origine des effets d'image que produisent certaines configurations linguistiques, dont, au premier chef, la métaphore), la fonction discursive (sollicitée dans l'argumentation par analogie par exemple, mais aussi dans les analogies que mettent en œuvre les discours didactiques ou de vulgarisation) et la fonction paraphrastique (qui assure les paraphrases intra-linguistiques et interlinguistiques [traduction¹⁹]).

Compte tenu de la nature cognitive du processus analogique, nous pouvons d'emblée récuser toute approche qui la limiterait à un processus conscient. La caractérisation saussurienne de l'analogie comme *procédé* par opposition au *processus* illustré par l'amalgame ne saurait donc être retenue. L'analogie est parfois un procédé, mais elle est le plus souvent un processus (non-conscient).

La fonction qui nous intéressera au premier chef pour aborder la question de la composition est bien sûr la fonction régularisatrice. Pour se convaincre qu'il s'agit bien d'un processus cognitif réel et pas seulement d'un artefact descriptif produit par le linguiste, il suffit de songer à quelques exemples bien connus du langage enfantin.

De nombreux enfants français, jusqu'à un certain âge, utilisent le composé *doigts de pieds* plutôt que *orteils* ou *trous de nez* plutôt que *narines*. Pourquoi? Dans l'hypothèse, probable, où les expressions n'ont pas été entendues par l'enfant avant qu'il les emploie, ces créations spontanées répondent à un besoin lexical par un moyen très économique au plan cognitif: des mots très usuels sont combinés selon une structure phraséologique (Nom + de + Nom) qui est aussi très fréquente.

¹⁸ Cf. Givón (1985): «Toutes choses égales par ailleurs, une expérience codée est d'autant plus facile à stocker, à récupérer et à communiquer que le code est isomorphe à l'expérience».

¹⁹ Une bonne traduction n'est autre qu'un bon analogue du texte traduit.

La composition, comme procédé néologique, requiert donc en premier lieu la disponibilité mentale d'une structure phraséologique.

Or cette structure n'est pas une "règle de composition" mais la propriété structurelle commune à de nombreuses constructions idiomatiques mémorisées. Autrement dit, le composé *doigts de pieds* peut être considéré comme une production analogique à partir d'une source représentée par *salle de bains*, *pied de table*, *bâton de colle*, et de très nombreuses autres constructions du même type, cette analogie étant précisément fondée sur la similarité structurelle (ou syntaxique) de ces constructions (Nom + de + Nom).

La question qui se pose, dans cette perspective analogique, est donc en premier lieu celle des schémas privilégiés de construction²⁰.

À titre exploratoire, nous avons relevé une centaine de composés récents (108 occurrences), repérés comme nouveaux dans trois dictionnaires (*Petit Larousse* 2017 et 2018; *Petit Robert* 2017)²¹. L'ensemble des unités phraséologiques est présenté dans le tableau 2 ci-dessous:

²⁰ Cette perspective est très proche de celle de la "morphologie naturelle", théorie des "préférences" qui «cherche à expliquer les préférences ou asymétries récurrentes de structure dans les langues du monde (et dans l'acquisition, la diachronie, etc.) par des motivations extérieures au langage» (Kilani-Schoch, Dressler, 2005: 17). Cependant, la morphologie naturelle ne s'intéresse guère à la composition dans la mesure où sa définition de la morphologie est centrée sur le mot et par conséquent sur les phénomènes d'inflexion et de dérivation qui supposent une base morphologique unique. Comme l'écrivent Dressler *et al.* (1987, p. 4) pour justifier la délimitation de leur propos dans l'ouvrage *Leitmotifs in natural morphology*: «compounding is not considered because a compound word must contain at least two base morphemes (roots ou stems)».

²¹ Ce relevé a été effectué à partir du site du Club d'orthographe de Grenoble, qui répertorie chaque année les mots nouveaux du *Petit Larousse* et du *Petit Robert*, et donne également un choix de composés nouveaux apparaissant dans les articles de ces dictionnaires (<https://orthogrenoble.net/mots-nouveaux-dictionnaires>). Nous avons relevé l'intégralité des composés nouveaux sélectionnés par le site pour les années 2017 et 2018. Le fait que cette liste soit issue d'une sélection effectuée non par des linguistes mais par des amateurs éclairés ne nous semble pas poser problème. Bien au contraire, elle ne peut que refléter plus fidèlement le "sentiment néologique" des locuteurs du français.

TABLEAU 2

Structure de quelques composés néologiques récents

Unités phraséologiques	Constructions	Unités phraséologiques	Constructions
*alignement des astres ou des planètes	NdeN	*voiture-bélier	NN
*amygdale cérébelleuse	NA	*vote d'adhésion	NdeN
*amygdale cérébrale	NA	*zone de confort	NdeN
*aplasie (médullaire)	NA	ancien *français	AN
*auditeur libre	NA	balcon *filant	Npp
*batterie-fanfare	NN	bureau *centralisateur	NA
*biscuit (rose) de Reims	NdeN	café, pain *suspendu	Npp
*cavalier législatif	NA	Cote *Argus, ou Argus	NN
*chèque emploi service universel (CESU)	Sigle	coup de *menton	NdeN
*colopathie fonctionnelle	NA	cuisine *fusion	NN
*complétion (automatique)	NA	culture *urbaine	NA
*conseil départemental	NA	droit *souple	NA
*consentement éclairé	Npp	droit à l'*oubli	NprépN
*continent de plastique	NdeN	droit de *réutilisation	NprépN
*couloir de nage	NdeN	économie *industrielle	NA
*coupe menstruelle	NA	élections *départementales	NA
*débit de kalach ou de mitraillette	NdeN	emploi d'*avenir	NdeN
*diagnostic différentiel	NA	épisode *cévenol	NA
*droit souple	NA	étude d'*impact environnemental	NdeNA
*échange colombien	NA	fibre *technique	NA
*économie collaborative	NA	film, cinéma d'*auteur	NdeN
*économie de partage	NdeN	fonds *vautour	NN
*encéphalogramme plat	NA	gâteau à la *broche	NprépN
*entreprise solidaire d'utilité sociale (ESUS)	NA	*trêve hivernale	NA
*État (fédéré)	Sigle	génération *boomerang	NN
*expérience de pensée	Npp	grand *djihad	AN
*expérimentation législative	NdeN	*djihad majeur	NA
*français moderne	NA	grande *prêtresse	AN
*frelon asiatique	NA	la *ligne claire	NA
*glace noire	NA	la *matière noire	NA
*habilitation familiale	NA	le *roi du pétrole	NdeN
*juge de ligne	NdeN	le fils, la fille du *facteur	NdeN
*loup solitaire	NA	loi de *simplification	NdeN
*marqueur social	NA	mélange des *genres	NdeN
*métaux lourds	NA	œil *directeur	NA
*paiement sans contact	NprépN	Parti *Pirate	NN
*pain au lait	NprépN	pause *mériidienne	NA
*palet breton	NA	personne *vulnérable	NA
		petit *djihad	AN
		*djihad mineur	NA
		pleine *conscience	AN

(segue)

TABLEAU 2 (*seguito*)

Unités phraséologiques	Constructions	Unités phraséologiques	Constructions
*permis (de conduire)	NdeInf	point d’*équivalence	NdeN
*prestation compensatoire	NA	pomme *paillasson	NN
*principe de participation du public	NdeNdeN	pratique *abusive	NA
*puits canadien	NA	préjudice *écologique	NA
*puits provençal	NA	principe d’*équivalence	NdeN
*réseau d’éducation prioritaire (REP)	Sigle	retard à l’*allumage	NprépDétN
*scène ouverte	NA	salle de *consommation	NdeN
*sclérose latérale amyotrophique (SLA)	Sigle	sport *électronique	NA
*soustraction de mineur	NdeN	suites de *couches	NdeN
*suite parentale	NA	sursaut *gamma	NN
*témoin (de connexion)	NdeN	syndrome de l’intestin	NdeDetN
*température ressentie	Npp	syndrome du côlon	NdeDetNA
*théorie du genre	NdeN	*irritable	
*transaction pénale	NA	tissu *technique	NA
		vêtement *technique	NA
		trouble déficitaire de l’*attention (TDA)	Sigle
		vin de *garde	NdeN
		zone *grise	NA

La répartition des types de structures est la suivante:

TABLEAU 3
Tendances des structures de composés néologiques

[Total = 108]	Nombre d’occurrences	Pourcentage
NA (Nom + Adj.)	50	46,3
NdeN (Nom + de + Nom)	26	24
NN (Nom + Nom)	9	8,3
AN (Adj. + N)	5	4,6
Sigle	5	4,6
NPrépN (Nom + préposition autre que de + Nom)	5	4,6
Npp (N+participe)	5	4,6
Autre	3	2,8

Ce corpus restreint fait apparaître une tendance nette à la préférence pour les structures [Nom + Adjectif] et [Nom + de + Nom]. Il est très probable que cette tendance se vérifierait pour un corpus plus étendu.

Au plan analogique, cela signifie que ces deux structures sont des sources très aisément mobilisables pour la création de composés. Mais comment expliquer cette préférence?

Une première explication réside dans la préférence globale pour les structures binaires, bien reconnue par la morphologie naturelle (Kilani-Schoch, Dressler, 2005, p. 17). Mais dans ce cas, il reste à expliquer pourquoi les structures [Nom + Adjectif] sont très nettement préférées aux structures [Adjectif + Nom] ou aux structures [Nom + Nom], qui sont elles aussi binaires mais beaucoup moins fréquentes dans le corpus.

Il convient ici de faire intervenir une autre source analogique, qui repose sur une caractéristique bien connue du français: la préférence pour l'ordre déterminé-déterminant²². La structure [Nom + Adjectif] respecte cet ordre (ce qui signifie qu'elle est analogue aux autres structures qui le respectent également), ce qui n'est pas le cas des deux autres structures concurrentes. Bien sûr, il ne s'agit que de tendances. Et pour chacune des occurrences du tableau 2, on pourrait s'attacher à examiner les raisons qui peuvent expliquer le choix de telle ou telle structure de composé.

Quant à la dimension "agglutinative" du mot composé, elle met en jeu un autre type d'analogie, hétérogène cette fois-ci, entre le plan du signe et celui du concept. C'est donc ici un aspect de la fonction iconique de l'analogie qui est mobilisée.

En effet, aussi variable qu'il soit, le figement qui caractérise les noms composés ou les unités phraséologiques peut être vu comme la manifestation d'une unité au plan conceptuel. Le mot composé présente une unité au plan du signe, qui se manifeste par l'impossibilité ou la difficulté de certaines transformations²³ (unité iconique de l'unité conceptuelle de la dénomination Kleiber, 2001, p. 36 et Monneret, 2017).

²² Voir par exemple Benveniste (1966, p. 91), Puvion et Sablayrolles (2003, p. 105).

²³ P. ex.: *un auditeur libre* > *un auditeur très libre*, *la liberté de l'auditeur*, *un auditeur qui est libre*, etc.: ces transformations sont impossibles sans altération du sens du nom composé.

Cependant, la tendance générale que connaissent de nombreuses langues à ce que Benveniste nomme la *synapse*, c'est-à-dire la tendance au regroupement et à l'unification d'éléments contigus, ne doit rien à l'analogie en elle-même.

Cette tendance résulte d'un autre processus fondamental, que l'on peut nommer la *fusion*, et qui se combine à l'analogie pour la production d'unités phraséologiques. La fusion est à la contiguïté ce que l'analogie est à la similarité: si l'analogie est un processus qui tend à l'*identification* d'éléments ou de structures en relation de similarité, la fusion est un processus qui tend à l'*unification* d'éléments ou de structures en relation de contiguïté. La théorie des deux axes de Jakobson gagnerait donc sans doute à être repensée sur un plan cognitif.

Bibliographie

- Bally C. (1951 [1909]), *Traité de stylistique française*, I et II, Librairie Georg & Cie et Librairie C. Klincksieck, Genève-Paris.
- Benveniste É. (1966 et 1974), *Problèmes de linguistique générale*, tomes I et II, PUF, Paris.
- Benveniste É. (1966), *Bulletin de la Société de Linguistique*, Paris, pp. 88-106.
- Bloomfield L. (1970 [1933]), *Le langage*, Payot, Paris.
- Clement C. A., Gentner D. (1991), *Systematicity as a selection constraint in analogical mapping*, in "Cognitive Science", 15, pp. 89-132.
- Dressler W., Mayerthaler W., Panagl O., Wurzel W. (1987), *Leitmotifs in Natural Morphology*, John Benjamins, Amsterdam.
- Evans R. (2000), *A comparison of rule-based and machine learning methods for identifying non-nominal it*, in D. N. Christodoulakis (ed.), *Proceedings of natural language processing – NLP 2000*, University of Patras, Greece, 2-4 June 2000, Springer Verlag, Berlin, pp. 233-41.
- Fischer O. (2000), *Grammaticalisation: Unidirectional, non-reversible? The case of to before infinitive in English*, in O. Fischer, A. Rosenbach, D. Stein (eds.), *Pathways of change: Grammaticalization in English*, John Benjamins, Amsterdam, pp. 149-69.
- Gentner D. (1983), *Structure-mapping: A theoretical framework for analogy*, in "Cognitive Science", 7/2, 155-170.
- Gentner D., Holyoak K. J., Kokinov B. (eds.) (2001), *The analogical mind: Perspectives from cognitive science*, The MIT Press, Cambridge-London.
- Gentner D., Markman A. B. (1995), *Similarity is like analogy*, in C. Cacciari (ed.), *Similarity*, Brepols, Bruxelles, pp. 111-48.
- Gentner D., Markman A. B. (1997), *Structure-mapping in analogy and similarity*, in "American Psychologist", 52, pp. 45-56.
- Gentner D., Smith L. (2012), *Analogical reasoning*, in V. S. Ramachandran (ed.), *Encyclopedia of human behavior*, Elsevier, Oxford (II ed.), pp. 130-6.

- Givón T. (1985), *Iconicity, isomorphism and non-arbitrary coding in syntax*, in J. Haiman (ed.), *Iconicity in syntax*, John Benjamins, Amsterdam.
- Hofstadter D. (2001), *Analogy as the core of cognition*, in D. Gentner, K. J. Holyoak, B. N. Kokinov (eds.), *The analogical mind: Perspectives from cognitive science*, The MIT Press, Cambridge-London, pp. 116-44.
- Hofstadter D., Sander E. (2013), *L'Analogie, cœur de la pensée*, O. Jacob, Paris.
- Hofstadter D., the Fluid Analogies Research Group (1995), *Fluid concepts and creative analogies. Computer models of the fundamental mechanisms of thought*, Basic Books, New York.
- Holyoak K. J., Thagard P. (1989), *Analogical mapping by constraint satisfaction*, in "Cognitive Science", 13, pp. 295-355.
- Holyoak K. J., Thagard P. (1995), *Mental leaps: Analogy in creative thought*, The MIT Press, Cambridge.
- Itkonen E. (2005), *Analogy as Structure and process, approaches in linguistics, cognitive psychology and philosophy of science*, John Benjamins, Amsterdam.
- Jakobson R. (1965), *À la recherche de l'essence du langage*, in "Diogenès", 51.
- Kilani-Schoch M., Dressler W. (2005), *Morphologie naturelle et flexion du verbe français*, NarrVerlag, Tübingen.
- Kleiber G. (2001), *Remarques sur la dénomination*, in "Cahiers de praxématique", 36, Presses universitaires de la Méditerranée, Montpellier, pp. 21-41.
- Markman A. B., Gentner D. (1993), *Structural alignment during similarity judgements*, in "Cognitive Psychology", 25, pp. 431-67.
- Monneret Ph. (2004), *Essais de linguistique analogique*, ABELL, Dijon.
- Monneret Ph. (2014), *L'iconicité comme problème analogique*, in "Le français moderne", 1, pp. 46-77.
- Monneret Ph. (2017), *Dénomination et analogie*, in G. Petit, P. Haillet, X.-L. Salvador (eds.), *La dénomination: lexique et discours*, Honoré Champion, Paris, pp. 95-114.
- Monneret Ph. (à paraître a), *Dénomination et analogie*, in *La dénomination: approches linguistiques et terminologiques*, Bibliothèque de l'Information Grammaticale Peeters, Petit Gérard, Leuven.
- Monneret Ph. (à paraître b), *Principes de linguistique analogique*, in Ph. Monneret, S. Mejri (éds.), *Langage et analogie*, 1, Éditions Lambert-Lucas, Paris.
- Mortureux M.-F. (2013 [2004]), *La lexicologie. Entre langue et discours. Deuxième édition revue et actualisée*, Armand Colin, Paris.
- Piaget J. (1975), *L'équilibration des structures cognitives: problème central du développement*, PUF, Paris.
- Polguère A. (2008), *Lexicologie et sémantique lexicale. Notions fondamentales*, Les Presses de l'Université de Montréal, Montréal.
- Pruvost J., Sablayrolles J. F. (2003), *Les néologismes*, PUF, Paris.
- Saussure F. de (1967 [1916]), *Cours de linguistique générale, édition critique préparée par Tullio de Mauro*, Éditions Payot & Rivages, Paris.

Dictionnaires

- Larousse P. (2016), *Le Petit Larousse illustré 2017*, Éditions Larousse, Paris.
Larousse P. (2017), *Le Petit Larousse illustré 2018*, Éditions Larousse, Paris.
Rey A. (2016), *Le Petit Robert de la langue française 2017*, Le Robert, Paris.

Sitographie

- Centre National de Ressources Textuelles et Lexicales (2017), *Refuir, verbe intransitif*, in <http://www.cnrtl.fr/definition/refuir>.
Centre National de Ressources Textuelles et Lexicales (2017), *Raller, verbe*, in <http://www.cnrtl.fr/definition/dmf/raller>.
Club d'orthographe de Grenoble (2017), *Mots nouveaux des dictionnaires, Corpus DiCo (Dictionnaires comparés)*, in <https://orthogrenoble.net/mots-nouveaux-dictionnaires>.