

Généricité et (in)définitude. Une analyse dans la théorie de l'optimalité

Henriëtte de Swart et Donka Farkas

Introduction

Depuis la thèse de Greg Carlson, publiée en 1980, les pluriels nus ont été le point de départ d'un grand nombre d'études portant sur la généricité. Cependant, les romanistes se rendent bien compte du fait que les pluriels nus sont soit pratiquement inexistants (en français d'aujourd'hui), soit inaptes à l'expression de la généricité (en italien, en espagnol, en roumain, etc.). On constate que, dans les langues romanes, la généralisation sur les individus ou la référence à l'espèce s'exprime principalement au moyen de pluriels définis. Cette observation soulève bien des questions. Cet emploi de l'article défini constitue-t-il une particularité des langues romanes? Y-a-t-il une différence de sens entre les pluriels nus de l'anglais et des langues germaniques d'une part et les pluriels définis des langues romanes (et autres) dans les contextes génériques? Qu'est-ce qui gouverne le choix de l'article utilisé dans les phrases génériques à travers les langues? Ces questions n'ont guère été examinées dans la littérature sur la généricité. Dans cet article, nous proposons une réponse dans le cadre de la théorie de l'optimalité.

1. Généricité définie et indéfinie

L'expression de la généricité dans les langues naturelles et l'interprétation des phrases génériques sont le sujet d'une vaste littérature linguistique et philosophique. Dans cet article, il nous est impossible de revoir cette littérature très riche, et d'en présenter une vue d'ensemble. Nous allons donc nous limiter à un problème très particulier, à savoir la différence entre les langues dans l'emploi des formes définies et indéfinies pour les SN pluriels à interprétation générique. Afin d'examiner cette question, nous étudions des langues qui présentent une opposition singulier/pluriel marquée dans la flexion nominale, et qui disposent d'articles définis et indéfinis. Les langues qui utilisent principalement les noms sans articles (singuliers et pluriels 'nus' comme dans les langues slaves, en chinois et en hindi, avec des interprétations définies, indéfinies et génériques, voir Chierchia 1998, Dayal 2004) se trouvent donc en dehors de notre typologie. Parmi les langues qui nous concernent se trouvent l'anglais (et d'autres langues germaniques) d'une part, et le français, d'autres langues romanes, le grec et le hongrois d'autre part. L'emploi de l'article défini est soumis à des règles très strictes, et ne varie pas essentiellement parmi ces langues dans les contextes épisodiques comme (1) et (2):

- | | | | |
|-----|----|-------------------------|------------|
| (1) | a. | The moon is round. | [anglais] |
| | b. | La lune est ronde. | [français] |
| | c. | A hold kerek. | [hongrois] |
| (2) | a. | The stars are shining. | [anglais] |
| | b. | Les étoiles brillent. | [français] |
| | c. | A csillagok csillognak. | [hongrois] |

On pourrait donner des paradigmes comparables pour les indéfinis singuliers et pluriels. La seule différence ici consiste dans la présence d'un article en français

moderne (l'article partitif *des*) versus l'absence d'article, et l'emploi de pluriels 'nus' dans les autres langues romanes, en hongrois, et en anglais. Cette différence n'affecte pas notre analyse, et nous traiterons les deux comme des indéfinis pluriels, suivant les approches de Bosveld-de Smet (1998) et de Roodenburg (2004b).

Dans les contextes génériques, le défini et l'indéfini singulier se comportent de la même manière à travers les langues, comme le montrent (3) et (4):

- | | | | |
|-----|----|---|------------|
| (3) | a. | The dinosaur is extinct. | [anglais] |
| | b. | Le dinosaure a disparu. | [français] |
| | c. | A dinoszaurusz kihalt. | [hongrois] |
| (4) | a. | A dog is dangerous when it is hungry. | [anglais] |
| | b. | Un chien est dangereux quand il a faim. | [français] |
| | c. | Egy kutya veszélyes mikor éhes. | [hongrois] |

Selon Beyssade (ce volume), le SN défini singulier en (3b) dénote l'espèce comme une unité atomique du domaine. Farkas et de Swart (2005) étendent cette analyse aux SN définis anglais et hongrois. Il n'y a donc pas de différence à travers les langues germaniques/romanes dans les environnements tels que (3). La généralisation générique sur les individus s'exprime avec un article indéfini au singulier dans toutes les langues en question. L'unicité associée au défini singulier bloque son emploi comme domaine de quantification, selon une règle pragmatique proposée par de Swart (1991, 1996). Parmi les singuliers, l'article indéfini est le seul à créer un domaine de quantification ouvert, intensionnel, selon Farkas et de Swart (2005). Les ressemblances dans les paradigmes (1)-(4) suggèrent fortement que les mêmes principes d'interprétation des articles définis et indéfinis sont actifs dans toutes les langues qui connaissent cette distinction. Cependant, il est bien connu que des différences existent dans l'expression de la généralité au pluriel:

- | | | | |
|-----|----|----------------------------|------------|
| (5) | a. | Dinosaurs are extinct. | [anglais] |
| | b. | Les dinosaurs ont disparu. | [français] |
| | c. | A dinoszauruszok kihaltak. | [hongrois] |

Comme les exemples en (5) l'illustrent, l'anglais (et les langues germaniques en général) utilisent des pluriels 'nus' pour renvoyer à l'espèce, tandis que le hongrois, le français et les langues romanes en général utilisent des pluriels définis dans ce contexte. C'est le contraste dans (5) qui nous occupe dans cet article. Par rapport à ce contraste, nous soulevons trois questions principales: (i) pourquoi le contraste dans la forme (article défini versus indéfini) pose-t-il problème aux analyses existantes de la généralité? (ii) comment ce contraste s'intègre-t-il dans une théorie générale et unifiée de l'interprétation des articles définis et indéfinis à travers les langues étudiées, basée sur les paradigmes (1)-(4)? (iii) comment ce contraste permet-il une interprétation des phrases (5a,b,c) en termes de référence à l'espèce, malgré la différence de forme? Nous examinons la question (i) dans la section 2, (ii) dans les sections 3 et 4, (iii) dans la section 5. La section 6 traite de quelques prédictions que fait notre analyse. La section 7 conclut.

2. Référence à l'espèce et généralisation générique

Selon Carlson (1980), les pluriels nus de l'anglais renvoient à l'espèce, parce qu'ils sont compatibles avec les prédicats du type *être en voie de disparition*, qui ne

s'appliquent qu'aux espèces (6a). Ils peuvent également exprimer une généralisation sur des individus dans le contexte d'un prédicat dit *individual-level* ou *non-spécifiant* (Kleiber 1981), qui dénote une propriété stable, sinon permanente, des individus. (6b). Finalement, la combinaison avec un prédicat dit *stage-level* ou *spécifiant* tel que *jouer dans la rue* donne lieu à une interprétation existentielle (6c).¹

- (6) a. Dinosaurs are extinct.
dinosaurés ont disparu
'Les dinosaurés ont disparu.'
- b. Dogs are intelligent.
chiens sont intelligents
'Les chiens sont intelligents.'
- c. Children were playing in the street.
enfants jouaient dans la rue
'Des enfants jouaient dans la rue.'

Carlson propose une analyse unifiée des trois cas exemplifiés dans (6). La référence à l'espèce offre une explication naturelle de l'acceptabilité de (6a). (6b) implique un opérateur qui lie l'espèce aux objets qui réalisent l'espèce. Finalement, la localisation spatio-temporelle des réalisations de l'espèce dans (6c) donne lieu à une interprétation existentielle.

Les indéfinis singuliers de l'anglais sont également compatibles avec les prédicats spécifiants et non-spécifiants, comme on le voit dans (7b), (7c). Cependant, ils n'acceptent pas la référence à l'espèce. (7a) est inacceptable, sauf dans une lecture taxonomique:

- (7) a. #A dinosaur is extinct.
'Un dinosauré a disparu.'
- b. A dog is intelligent.
'Un chien est intelligent.'
- c. A dog was playing in the street.
'Un chien jouait dans la rue.'

Etant donné que les indéfinis singuliers ne renvoient pas à l'espèce, Carlson propose une analyse tout à fait différente des pluriels nus et des indéfinis singuliers.

La notion de quantification non-sélective offre une analyse différente des exemples (6) et (7). Au lieu de mettre l'accent sur une analyse unifiée du pluriel nu, l'analyse de la généricité dans le cadre de la Théorie de la Représentation du Discours (Discourse Representation Theory: DRT) a souligné la parité de (6b)/(7b) et de (6c)/(7c). Dans les deux cas, on analyse l'indéfini comme une expression introduisant une variable libre. Cette variable libre est liée par la clôture existentielle dans les exemples (6c, 7c), et par un quantificateur générique, quasi-universel dans les exemples (6b, 7b) (Farkas and Sugioka 1983, Diesing 1992, Kratzer 1995, Krifka et al. 1995). La principale différence entre l'indéfini singulier et le pluriel nu consiste dans le fait que le pluriel nu présente une ambiguïté absente pour l'indéfini singulier.

¹ La distinction établie par Carlson entre prédicats *individual-level* et *stage-level* n'est pas identique à la distinction établie par Kleiber entre prédicats non-spécifiants et spécifiants, mais en gros, les deux classifications se recouvrent. Les différences ne sont pas pertinentes pour notre article, mais cf. Dobrovie-Sorin (1997) pour une discussion plus approfondie.

La quantification non-sélective et la référence à l'espèce constituent ainsi deux interprétations du pluriel nu indépendantes l'une de l'autre.

Aucune des deux analyses ne s'applique très bien aux langues romanes. Dans les langues romanes, le pluriel nu est soit pratiquement inexistant (c'est le cas du français moderne), soit pratiquement absent des contextes génériques (dans les autres langues romanes). Longobardi (2001) donne les exemples suivants pour illustrer l'incompatibilité des pluriels nus de l'italien avec la référence à l'espèce et les prédicats non-spécifiants:

- (8) a. *Elefanti di colore bianco sono estinti. [italien]
 éléphants de couleur blanche ont disparu
 b. *Ucelli di zone paludose sono intelligenti.
 oiseaux de régions marécageuses sont intelligents
 c. Elefanti di colore bianco hanno creato in passato una grande curiosità.
 éléphants de couleur blanche ont suscité dans le passé une grande curiosité
 'Des éléphants de couleur blanche ont suscité une grande curiosité dans le passé.'

En principe, les pluriels nus sont compatibles avec les prédicats spécifiants seulement. (8c) montre qu'ils ont une interprétation existentielle dans ce cas-là. Les exemples de Longobardi vont à l'encontre des observations faites par Chierchia (1998), qui insiste sur l'occurrence des pluriels nus dans les contextes génériques. Nous verrons dans la section 6.1 que ses exemples concernent le phénomène qu'on appellera *pseudo-généricité*. En général, les contextes génériques demandent l'emploi de l'article défini en italien (exemples de Longobardi 2001) :

- (8') a. Gli elefanti di colore bianco sono estinti.
 'Les éléphants de couleur blanche ont disparu.'
 b. Gli ucelli di zone paludose sono intelligenti.
 'Les oiseaux de régions marécageuses sont intelligents.'

Les autres langues romanes suivent le modèle de l'italien (cf. Dobrovie-Sorin et Laca 2003). Nous illustrons avec le roumain :

- (9) a. Dinozaurii au dispărut. [roumain]
 dinosaure.Pl.Def . ont disparu
 'Les dinosaures ont disparu.'
 b. Cîinii sînt inteligenți.
 chien.Pl.Def sont intelligents
 'Les chiens sont intelligents.'
 c. Se joacă copii pe stradă.
 Refl. jouent enfant.Pl dans la rue
 'Des enfants jouent dans la rue.'

Les exemples (8') et (9) illustrent le fait que les langues romanes utilisent les pluriels définis pour exprimer la référence à l'espèce (8'a, 9a) et la généralisation générique (8'b, 9b). Les pluriels nus sont utilisés pour l'interprétation existentielle (8c, 9c). Le français occupe une position spéciale dans le système, parce qu'il n'emploie presque plus de noms nus (mises à part les constructions de coordination relevées par Roodenburg 2004). Le paradigme implique donc une comparaison entre définis et

indéfinis comportant l'article partitif *des*. *Modulo* cette modification, les observations sont tout à fait comparables (cf. également Dobrovie-Sorin et Laca 2003). Autrement dit, l'article partitif ne figure pas dans les phrases génériques (10a,b), mais se limite aux contextes existentiels (10c). L'expression de la généricité demande l'article défini (10')

- (10) a. *Des dinosaures ont disparu.
 b. *Des chiens sont intelligents.
 c. Des enfants jouent dans la rue.
- (10') a. Les dinosaures ont disparu.
 b. Les chiens sont intelligents.

L'analyse unifiée des pluriels nus proposées par Carlson ne s'applique pas très bien aux langues romanes, parce qu'il n'y a pas unité de forme entre les séries a, b et c. On pourrait essayer de sauver l'analyse basée sur les espèces en considérant les définis des langues romanes comme des expressions qui renvoient à l'espèce (cf. Longobardi 1994, 2001, Dobrovie-Sorin et Laca 1999), ce qui rendrait compte des exemples du type (8'a, 9a, 10'a). Cependant, cela n'expliquerait pas l'emploi épisodique des définis dans les phrases telles que (11):

- (11) a. Copiii se joacă pe stradă. [roumain]
 enfant.Pl.Def. Refl. jouent sur rue.
 b. Les enfants jouent dans la rue. [français]

Evidemment, on pourrait donner des exemples comparables pour l'anglais et pour d'autres langues germaniques:

- (12) a. The children are playing in the street [anglais]
 'Les enfants jouent dans la rue.'
 b. De kinderen spelen op straat. [néerlandais]
 'Les enfants jouent dans la rue.'

On constate que l'opposition entre article défini et indéfini dans les contextes épisodiques est la même dans les langues romanes et dans les langues germaniques. Cette observation implique qu'on ne peut pas adopter l'analyse unifiée de Carlson pour les langues romanes sans faire des concessions considérables.²

Si on essaie d'appliquer l'analyse de la quantification non-sélective aux langues romanes, on se heurte à un problème comparable. Pour expliquer les exemples du type (8'b, 9b), il faudrait supposer que les définis introduisent des variables qui peuvent être liées par un quantificateur générique. Vergnaud et Zubizarreta (1992) défendent l'idée que l'article défini dans les contextes génériques est à considérer comme un explétif qui ne contribue rien à la sémantique de la phrase générique. Krifka et al. (1995) parlent d'un marqueur 'thématique'. Si l'article défini constitue en effet un explétif ou un marqueur thématique, on pourrait maintenir une

² Longobardi (1994, 2001) fait appel à la notion de paramétrisation syntaxique des expressions référentielles pour maintenir une version de l'analyse de Carlson. Il y a certainement des ressemblances entre ses propositions et nos idées. Cependant, nous pensons que l'approche dans la théorie de l'optimalité offre un cadre plus riche que la paramétrisation, surtout dans l'explication des effets de la pseudo-généricité et de la généricité anaphorique.

analyse en termes de quantification non-sélective par un quantificateur générique. Cependant, on ne voit pas bien pourquoi l'article défini se prêterait à cet usage d'explétif ou de marqueur thématique. En plus, cette analyse n'explique ni la possibilité qu'ont les pluriels définis d'exprimer la référence à l'espèce (8'a, 9a, 10'a), ni l'emploi 'standard' de l'article défini dans les contextes épisodiques (11, 12), où l'article ne fonctionne pas comme explétif, ni le contraste entre les langues germaniques et les langues romanes.

On pourrait bien sûr dire que c'est un trait particulier des langues romanes que d'exprimer la généralité au moyen d'un article défini, plutôt que d'utiliser le pluriel indéfini (en français) ou le pluriel nu (en italien, en roumain, etc.). Et on pourrait en rester là. Mais à notre avis, ce serait méconnaître le problème dans la mesure où d'autres langues illustrent le même paradigme, sans être liées aux langues romanes. Longobardi (2001) mentionne le grec comme un cas pertinent, et donne les exemples sous (13):

- (13) a. *Asproi elephantes echoun exaphanisthei.
 blancs éléphants ont disparu.
 b. Oi asproi elephantes echoun exafanisthei.
 'Les blancs éléphants ont disparu.'

Le hongrois suit également le modèle des langues romanes :

- (14) a. A dinoszauruszok eltűntek. [hongrois]
 Def dinosaurs.pl disparaître.Parf
 'Les dinosaures ont disparu.'
 b. A kutyák okosak.
 Def chien.pl intelligent.pl.
 'Les chiens sont intelligents.'
 c. Gyerekek játszanak a kertben.
 enfant.Pl. jouer.Prés.Pl Def jardin.dans
 'Des enfants jouent dans le jardin.'
 d. A gyerekek játszanak a kertben.
 Def enfant.pl. jouer.Prés.Pl Def jardin.dans
 'Les enfants jouent dans le jardin.'

Ces exemples montrent que le pluriel défini exprime la référence à l'espèce (14a) ou la généralisation générique (14b), tandis que le pluriel nu a une interprétation existentielle (14c). Les définis pluriels ont également un emploi épisodique régulier (14d). En principe, le fait que le grec, le hongrois et les langues romanes partagent le même paradigme pourrait être le fruit du hasard, mais il est plus probable que cette disposition des données révèle un problème plus général, et plus intéressant.

Il convient donc de regarder l'emploi générique de l'article défini au pluriel de plus près. Cependant, ni la théorie de Carlson, ni la quantification non-sélective ne se prête facilement à une analyse du problème, car les deux approches partent de l'anglais, et mettent l'accent sur le pluriel nu d'une part et l'indéfini singulier d'autre part. Afin de rendre compte des SN génériques définis pluriels des langues romanes et du hongrois, nous adopterons une analyse de l'opérateur générique en termes de quantification sur des situations ou des éventualités (de Swart 1991, 1996 et d'autres). Cette approche nous permettra de traiter les pluriels nus de l'anglais comme des indéfinis qui impliquent un quantificateur existentiel, et les pluriels définis comme des

expressions de la référence déterminée dans le sens de Farkas (2002). Dans ce cadre, nous pourrions démontrer que les définis et les indéfinis sont également appropriés à l'expression de la généricité. Cette analyse sera développée dans la section 5.

Mais avant de présenter le détail de notre analyse, il convient de se poser la question du choix entre article défini et indéfini. Pourquoi l'anglais et d'autres langues germaniques utilisent-ils les pluriels nus pour exprimer la généricité, tandis que les langues romanes et le hongrois utilisent les pluriels définis dans ce contexte? La question est d'autant plus incontournable qu'on n'observe pas de grandes différences entre les deux groupes de langues en ce qui concerne l'emploi de l'article défini dans les contextes épisodiques (voir ci-dessus). Robertson (2004) a formulé un début de réponse à cette question en soutenant l'hypothèse que les pluriels définis du français n'ont pas les mêmes présuppositions de familiarité que leurs expressions correspondantes en anglais, ce qui permettrait leur emploi générique. L'analyse de Robertson est incomplète dans la mesure où elle n'explique pas pourquoi le caractère présuppositionnel plus relâché des définis français n'est pas visible dans les contextes épisodiques du type de (11), (12) ou (15). Elle n'explique pas non plus d'où vient cette différence entre les deux langues. Dans cet article, nous proposons une analyse plus approfondie, qui rend compte des divergences dans les contextes génériques aussi bien que des convergences dans les contextes épisodiques, et qui fonde les deux cas sur le contraste général entre article défini et indéfini. Nous prenons notre point de départ dans l'analyse du défini développée par Farkas (2002) (section 3). Farkas suppose que le SN défini indique la référence déterminée du référent. Cette notion implique deux caractéristiques centrales des définis: la référence unique (ou maximale) et la familiarité. Les emplois prototypiques de l'article défini dans les contextes épisodiques satisfont à la maximalité autant qu'à la familiarité. Nous envisageons les contextes génériques comme un cas hybride, où la référence maximale et la familiarité entrent en conflit. Le cadre de la théorie de l'optimalité nous permettra de résoudre ce conflit. Dans la section 4 nous adopterons une hiérarchie différente des contraintes pour les langues germaniques d'un côté, et pour les langues romanes, le grec et le hongrois de l'autre.

3. La référence déterminée: maximalité et familiarité

Nous adoptons l'analyse des définis développée par Farkas (2002). Selon Farkas, les SN définis et indéfinis sont à traiter comme une paire d'expressions marquée/non-marquée. C'est-à-dire que la notion de 'référence déterminée' caractérise les définis, tandis que les indéfinis ne sont tout simplement pas caractérisés par cette propriété (cf. Hawkins 1991). La référence déterminée implique l'absence de choix pour la valeur du référent dans le modèle. Elle subsume la notion de référence unique (pour les singuliers, 15a) ou de référence maximale (référence à la somme maximale d'individus pour les noms non-comptables et les pluriels, 15b), et la notion dynamique de familiarité (15c,d):³

- (15) a. Le soleil brille.
b. Les planètes tournent autour du soleil.
c. J'ai vu Marie ce matin. La pauvre femme avait l'air malade.
d. J'ai vu Paulette et Jules ce matin. Les deux gosses s'amusaient bien dans le parc.

³ Pour des raisons de temps et d'espace, nous laissons les noms non-comptables de côté dans cet article.

L'analyse de Farkas est formulée dans le cadre de la Théorie des Représentations du Discours (DRT, abréviation de l'anglais *Discourse Representation Theory*) de Kamp et Reyle (1993).⁴ Dans ce cadre théorique, les SN indéfinis, tout comme les SN définis, introduisent des référents de discours (rd). Nous supposons donc avec Farkas que l'article défini introduit un rd x dont la référence est déterminée (ce qui est indiqué par la notation $!x$) par rapport à la représentation du discours (DRS) construite jusque-là. Tous les rd introduits par les SN définis de (15) ont la propriété de référence déterminée, car la valeur de x dans le modèle est fixe: il n'y a pas de choix.

Comme cet article porte sur les pluriels, il faut préciser la notion de maximalité impliquée dans les exemples (15b) et (15d). Nous suivons ici les analyses de Link (1983), d'Ojeda (1993), de Beyssade (ce volume) et d'autres, qui ont proposé de construire le modèle ontologique des noms non-comptables et des pluriels au moyen de la structure mathématique de treillis. Le treillis est défini comme un ensemble d'éléments ordonnés par une relation d'inclusion. Si on ajoute l'opération d'union, on obtient une structure où on peut construire la somme d'un certain nombre d'éléments atomiques et non-atomiques. Le SN défini dénote le supremum du treillis (dans le contexte), donc la somme de tous les éléments contenus dans ce treillis, et de leurs sommes. Dans le cas du singulier, le treillis ne contient qu'un seul élément atomique dans le contexte, donc la notion de maximalité se confond avec l'unicité. Dans le cas du pluriel, le maximum nous rend l'ensemble complet. Dans (15d), le défini anaphorique renvoie à la somme des gosses introduits dans le discours préalable, c'est-à-dire, Paulette \oplus Jules. Dans (15b), la somme des planètes peut être calculée sur la base de notre connaissance du monde, même si le rd n'a pas encore été introduit dans la DRS.

La caractérisation des définis en termes d'une combinaison de maximalité et de familiarité reflète l'intuition que l'interprétation des SN a un côté statique (référence) et un côté dynamique (mise à jour d'une information préalable). Nous proposons de traduire les deux parties du sens des définis sous forme de deux contraintes:

- ♣ FidélitéMax (FMax)
Réfléter les traits de maximalité de l'input dans l'output.
- ♣ *Def/-Fam
Eviter les définis non-familiers.

La contrainte FidélitéMax est une contrainte qui établit un lien entre input et output. Si on va des sens aux formes (perspective de la génération donc de la syntaxe), nous nous mettons dans la position du sujet parlant qui essaie de choisir la meilleure forme pour exprimer un certain sens. Si l'input est un rd marqué [+Max], la contrainte FidélitéMax nous demande d'utiliser un article défini. La contrainte *Def/-Fam bloque l'emploi de l'article défini pour les rd marqué avec le trait [-Fam].⁵ Par défaut, les rd 'nouveaux' requièrent donc l'article indéfini.

Dans la plupart des théories linguistiques, les règles utilisées sont de nature absolue, et n'admettent aucune exception. Dans la théorie de l'optimalité (OT) par

⁴ Voir Corblin (2002) pour une application au français.

⁵ Cette contrainte ressemble beaucoup à la contrainte 'Les définis sont familiers' proposée par Beaver (2004). Nous choisissons la formulation 'inverse' pour souligner le fait que l'emploi de l'article peut être bloqué par une violation de cette contrainte.

contre, les contraintes ne sont pas absolues, et peuvent être violées. Les cas bien connus du type ‘exception à la règle’ impliquent alors la transgression d’une contrainte assez générale, motivée par le désir de satisfaire à une contrainte spécifique plus importante. Dans la théorie OT on suppose que les contraintes sont de nature universelle, mais que la hiérarchie des contraintes, qui décide de l’importance relative des contraintes, peut varier d’une langue à l’autre. Dans cet article, nous déterminons une syntaxe OT, où l’input est un sens, et l’output une forme. Nous supposons que les deux contraintes formulées ci-dessus régissent le choix de l’article dans la production d’un SN. Dans un contexte épisodique où le rd est marqué par les traits [+Max], [+Fam], il est possible de satisfaire aux deux contraintes. Avec cette configuration des traits, l’article défini est de mise dans toutes les langues qui s’en servent. Par conséquent, il n’y a pas de différence entre les langues considérées ici, dans les contextes épisodiques. Dans les contextes génériques ‘standard’ ou prototypiques par contre, le rd est marqué par les traits [+Max], [-Fam]. Cette configuration implique un conflit entre les propriétés de maximalité et de familiarité. Le choix entre article défini ou article indéfini/pluriel nu dépend maintenant de l’importance relative des deux contraintes mentionnées. Dans la section 4, nous défendrons l’idée que les langues germaniques et les langues romanes divergent précisément sur ce point.

4. Choix de l’article: analyse dans la théorie de l’optimalité

Comme nous l’avons dit rapidement à la fin de la section 3, nous partons de l’idée que le rd impliqué dans la référence à l’espèce ou la généralisation générique est marqué par les traits sémantiques [+Max] et [-Fam]. Cette caractérisation mérite d’être motivée. Farkas (2002) traite de la notion de référence déterminée dans les contextes épisodiques. Dans ces contextes, on perçoit assez clairement quels sont les membres du treillis, et comment il faut construire la somme de ces éléments qui constitue le référent du SN défini. Maintenant il faut étendre cette notion aux contextes génériques. Nous suivons ici Ojeda (1993), qui propose d’identifier l’espèce avec le sommet du treillis, et qui permet au treillis d’impliquer des objets dans différents mondes possibles (voir également Chierchia 1998, Dayal 2004, Beyssade ce volume). Etant donné ce cadre d’analyse, on peut dire que dans les phrases du type (6a, 8’a, 9a, 10’a, 14a), on renvoie à l’espèce dans son ensemble. De même, dans les phrases du type (6b, 8’b, 9b, 10’b, 14b) on renvoie à l’ensemble des réalisations de l’espèce à travers le temps et à travers les mondes possibles. Le rd a donc assurément un caractère maximal.

En même temps, le rd générique n’est pas familier dans le contexte discursif: il ne s’agit ni d’un emploi anaphorique, ni d’un cas d’accommodation. La construction de la généralité au pluriel maintient l’accès aux individus constituant la classe, comme le souligne Beyssade (ce volume). Par définition, les membres de la classe ne sont pas tous connus, car ils existent dans des mondes possibles différents. Dans le cas standard, la nature ouverte de la classe est donc incompatible avec la familiarité du rd.

En termes de traits sémantiques, les rd génériques ont donc une configuration mixte. D’une part, le trait [+Max] devrait entraîner l’emploi de l’article défini, d’autre part le trait [-Fam] devrait bloquer l’emploi de l’article défini. Les différents paradigmes introduits dans la section 1 ci-dessus montrent que les langues résolvent ce conflit de deux façons différentes. La grammaire qui adopte la hiérarchie FMax >> *Def/-Fam donne la priorité à la contrainte FidélitéMax, et accepte si nécessaire une infraction de la contrainte *Def/-Fam. Ainsi, la syntaxe OT génère un article défini pour exprimer la généralité. La grammaire qui adopte la hiérarchie *Def/-Fam >> FMax donne plus d’importance à la contrainte *Def/-Fam, si nécessaire aux dépens de

la contrainte FMax. La syntaxe OT bloque donc l'emploi de l'article défini pour l'expression de la généralité, et retombe sur l'indéfini en tant que terme neutre, non marqué de la paire. Les deux hiérarchies *Def/-fam >> FMax et FMax >> *Def/-fam sont illustrées dans les tableaux 1 et 2, avec des données de l'anglais et du hongrois respectivement:

Tableau 1: production de SN génériques indéfinis (langue exemple: anglais)

Sens	Forme	*Def/-fam	FMax
Gen _x (Chien(x), Intellig(x)) [+Max] [-Fam]			
	☞ Dogs are intelligent Chiens sont intelligents		*
	The dogs are intelligent Les chiens sont intelligents.	*	

Tableau 2: production de SN génériques définis (langue exemple: hongrois)

Sens	Forme	FMax	*Def/-fam
Gen _x (Chien(x), Intellig(x)) [+Max] [-Fam]			
	Kutyák okosak. Chiens sont intelligents	*	
	☞ A kutyák okosak. Les chiens sont intelligents		*

Les deux tableaux établissent la relation entre le sens (input) et la forme (output). Le sens qu'on veut exprimer est représenté de façon informelle comme une généralisation générique sur les chiens, au moyen d'un quantificateur générique liant la variable x. Cette représentation sera affinée dans la section 5, mais elle suffit pour l'instant. La configuration [+Max] [-Fam] fixe les traits du rd dans les phrases exprimant une généralisation générique. L'ordre de gauche à droite dans les tableaux reflète la hiérarchie des contraintes: plus une contrainte est importante, plus elle se trouve à gauche, plus il faut éviter de la violer. Comme on le voit dans les deux tableaux, il est impossible de trouver une forme qui satisfasse aux deux contraintes: chaque forme implique la violation de l'une des deux contraintes, indiquée par un astérisque *. Le conflit est résolu en faveur d'une forme indéfinie dans le premier tableau et en faveur d'une forme définie dans le deuxième tableau. Cette solution suit la stratégie qui dit qu'on peut violer une contrainte plus faible, afin de satisfaire à une contrainte plus forte. Le poids différent des deux contraintes dans les deux grammaires, reflété dans l'ordre de gauche à droite, fait donc toute la différence. N'importe quelle langue germanique pourrait remplacer les exemples anglais du premier tableau. La grammaire de toutes ces langues adopte la hiérarchie *Def/-Fam >> Fmax. Elles génèrent donc des SN indéfinis, plutôt que des SN définis, pour exprimer une généralisation générique. Toutes les langues romanes, ainsi que le grec et le hongrois, adoptent la hiérarchie FMax >> *Def/-Fam dans leur grammaire, et emploient donc l'article défini dans les contextes génériques.

Les tableaux 1 et 2 montrent qu'une violation de l'une des deux contraintes est inévitable à cause de la configuration de traits sémantiques du rd générique. Notre analyse repose donc essentiellement sur l'architecture de la théorie de l'optimalité, qui

implique des contraintes violables, qui imposent des contraintes potentiellement opposées. L'ordre hiérarchique entre les contraintes nous permet de résoudre le conflit, et de déterminer l'expression optimale du sens générique. On dit souvent que 'le gagnant ramasse la mise' pour exprimer le fait que les candidats perdants ne sont pas 'moins bien', mais tout simplement 'mauvais.' Les formes sous-optimales n'émergent donc pas du tout. En d'autres termes, l'anglais et les autres langues germaniques bloquent l'article défini des contextes génériques standard, tandis que l'article défini est toujours de mise pour les langues romanes, le grec et le hongrois. Nous finissons donc par avoir deux classes de langues en ce qui concerne l'expression de la généralité, qui correspondent aux deux ordres hiérarchiques possibles entre les contraintes FMax et *Def/-Fam.

L'analyse proposée s'étend naturellement aux phrases exprimant la référence à l'espèce. Nous suivons Carlson (1980) en supposant que les prédicats du type *être en voie de disparition* demandent un sujet qui renvoie directement à l'espèce. Nous utilisons des rd portant l'indice 'k' pour distinguer les entités qui renvoient à l'espèce (*kind* en anglais) des entités individus. Evidemment, chaque rd du type x_k a le trait [+Max], car nous supposons que l'espèce plurielle dénote le sommet du treillis (voir également Chierchia 1998, Dayal 2004, Beyssade ce volume). Dans les contextes où l'espèce n'a pas été mentionnée dans le discours précédent, le rd x_k aura également le trait [-Fam], car les réalisations de l'espèce ne sont pas connues, étant donné le caractère intensionnel du treillis. L'interaction entre les deux contraintes pertinentes donne lieu aux tableaux suivants :

Tableau 3: référence à l'espèce (langue exemple: anglais)

Sens	Forme	*Def/-Fam	FMax
Dinosaure(x_k) & Disparu(x_k) [+Max] [-Fam]			
☞	Dinosaurs are extinct dinosaures ont disparu		*
	The dinosaurs are extinct Les dinosaures ont disparu	*	

Tableau 4: référence à l'espèce (langue exemple: hongrois)

Sens	Forme	FMax	*Def/-Fam
Dinosaure(x_k) & Disparu(x_k) [+Max] [-Fam]			
	Dinoszauruszok eltűntek dinosaures ont disparu	*	
☞	A dinoszauruszok eltűntek Les dinosaures ont disparu		*

La prédication directe de propriétés de l'espèce conduit à la même répartition des formes du SN dans les langues germaniques et dans les langues romanes, car la configuration des traits du rd ne change pas par rapport à l'expression de la généralisation générique. Comme c'est la configuration des traits de maximalité et de familiarité qui détermine le choix de l'article, les langues qui utilisent l'article défini pour exprimer les généralisations génériques feront de même pour la référence à

l'espèce. En d'autres termes, ce sont les propriétés référentielles et dynamiques du rd qui déterminent le choix de l'article, indépendamment de la structure générique (généralisation sur les réalisations de l'espèce, tableaux 1, 2 ou référence à l'espèce, tableaux 3, 4). Par conséquent, notre analyse fait la prédiction qu'il n'y a pas de divergence à l'intérieur d'une langue entre les deux types de phrases génériques. Une langue où on exprimerait la référence à l'espèce au moyen d'un pluriel défini, et la généralisation générique au moyen d'un pluriel indéfini serait donc un contre-exemple à notre analyse. Nous n'avons pas trouvé de langue de ce type. Nous laissons au lecteur la tâche de vérifier que les résultats donnés dans les tableaux 3 et 4 pour l'anglais et le hongrois peuvent être reproduits pour les autres langues germaniques, les langues romanes, et le grec. Notre analyse caractérise l'anglais et les langues germaniques comme des langues à 'haute familiarité', car elles donnent la priorité au blocage de l'article défini pour les rd non-familiers. Par contre, les langues romanes, le grec et le hongrois sont à considérer comme des langues à 'haute maximalité', car elles demandent que la nature maximale du rd soit reflétée dans la forme. Cette analyse offre une explication des intuitions de Robertson (2004), qui dit que les SN définis du français ont des présuppositions de familiarité plus relâchées que les SN définis de l'anglais.

La grammaire OT que nous avons développée rend compte de deux classes de langues, illustrées ici par l'anglais et les langues germaniques d'une part, et les langues romanes et le hongrois de l'autre. Sur la base de notre analyse, nous formulons la conjecture que toute langue qui a des articles définis et indéfinis suit l'un de ces deux modèles pour l'expression de la généricité au pluriel.

Maintenant que nous avons établi qu'un input sémantique identique conduit à deux formes distinctes dans deux classes de langues, et que le choix est déterminé par la grammaire OT, la question se pose de savoir si les deux formes permettent de récupérer le même sens. En d'autres termes, les SN génériques définis et indéfinis dans les langues germaniques et romanes ont-ils le même sens ou pas? Dans la section 5, nous mettons en place un système d'interprétation des SN définis et indéfinis qui accorde deux sens équivalents aux phrases génériques des deux classes de langues. Cette analyse ne sera pas formulée dans le cadre de la théorie de l'optimalité, car il n'est pas question de compétition entre plusieurs sens différents.

5. La sémantique des SN génériques définis et indéfinis

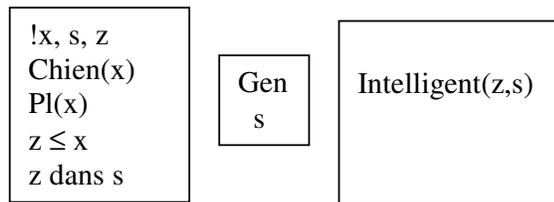
L'analyse des pluriels définis et indéfinis que nous avons proposée dans la section 2 ci-dessus part de l'idée qu'il s'agit d'une paire de formes: le défini est la forme marquée, l'indéfini est la forme non marquée. Le SN défini a la propriété de référence déterminée, ce qui veut dire que la valeur du rd dans le modèle n'offre pas de liberté de choix. Cette propriété de référence déterminée est absente de la sémantique du SN indéfini. Il faut souligner que cette analyse des SN définis et indéfinis est en principe indépendante du caractère épisodique ou générique de la phrase. Comme Farkas (2002) ne traite que d'exemples épisodiques, il faut étendre son analyse aux contextes génériques. De même, Farkas (2002) met l'accent sur les SN définis singuliers, tandis que cet article traite les SN pluriels. Il faut donc incorporer une sémantique de la morphologie du nombre dans notre système d'interprétation. Nous suivons ici la proposition de Farkas et de Swart (2003) selon laquelle la flexion plurielle nominale introduit la présupposition qu'il existe un rd qui porte le prédicat de pluralité. Cette présupposition peut être soit liée par le rd introduit par le déterminant (article indéfini, article défini ou quantificateur), soit « accommodée » (en l'absence d'un déterminant, à savoir dans le cas des pluriels 'nus'). Il s'agit ici de la notion d'accommodation

introduite par Lewis (1979), et interprétée dans le cadre de la DRT par van der Sandt (1992).⁶

En ce qui concerne l'analyse de la généralité, nous prenons pour point de départ de Swart (1991, 1996) et d'autres, qui ont proposé de traiter la généralisation générique au moyen d'un opérateur générique portant sur des situations ou des événements. Ainsi, l'opérateur générique ne lie pas directement les rd introduites par le SN défini ou indéfini, mais il les lie indirectement à travers les situations auxquelles ces rd participent. Nous suivons Ferreira (2005), qui propose d'interpréter l'opérateur générique comme un défini pluriel portant sur des éventualités. L'interprétation des phrases génériques du type *A kutyák okosak* (16) et *Dogs are intelligent* (17) peut ainsi être représentée par les DRS K_1 et K_2 :

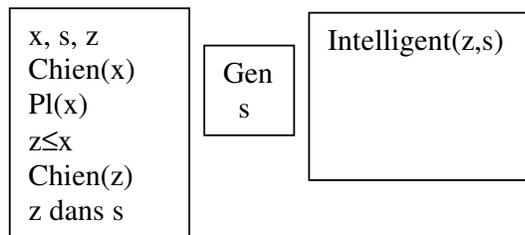
- (16) A kutyák okosak. [hongrois]
Les chiens sont intelligents.

DRS K_1



- (17) Dogs are intelligent. [anglais]
Chiens sont intelligents.

DRS K_2



La comparaison des deux DRS révèle qu'elles sont presque identiques. Dans les deux représentations, l'opérateur générique GEN introduit une opération ayant pour effet la subdivision de la DRS en deux boîtes subordonnées, liées par l'opérateur GEN qui généralise sur les situations s . La subdivision est sensible à la distinction entre *arrière-plan* et *focus*. Comme le prédicat constitue la partie focalisée de la phrase, il va dans la boîte de droite, tandis que l'information qui vient du sujet va dans la boîte de gauche, dans l'arrière-plan du prédicat. Le rd x introduit par le sujet est un

⁶ Comme l'explique Corblin (2002: 215), une expression présupposant une information antérieure qui a été employée sans que cette information figure explicitement dans le discours ne peut pas faire appel au processus du liage de la présupposition. Il faut alors l'accomoder dans le sens de Lewis (1979), c'est-à-dire faire comme si l'information présupposée était accessible dans ce contexte, et donc l'inscrire dans la représentation du discours.

participant de la situation s . Le rd dans la boîte gauche de K_1 est introduit par l'article défini *les*. Le déterminant lie le rd présupposé par la morphologie plurielle sur le nom *chiens*. L'absence de déterminant dans la phrase anglaise (17) entraîne un processus d'accommodation du rd présupposé par la morphologie plurielle sur le nom *dogs*, ce qui revient à postuler l'existence d'un individu correspondant dans la DRS. Les DRS ne sont donc pas construites de la même façon (voir Farkas et de Swart 2003 pour les détails de l'analyse), mais le résultat final a les mêmes conditions de vérité.

Quant à la structure quantificationnelle, l'opérateur GEN dans les DRS K_1 et K_2 porte sur toutes les situations s qui impliquent un z (où z est un chien, partie du rd x pluriel introduit par *chiens*) et les étend à une situation où z est intelligent. Etant donné qu'*être intelligent* est une propriété distributive (les individus sont intelligents en tant qu'individus plutôt qu'en tant que groupe), nous introduisons la condition $z \leq x$ pour avoir accès aux individus atomiques. Nous en resterons à un niveau relativement informel en ce qui concerne le phénomène de la distributivité, qui a été bien étudié ailleurs (cf. Verkuyl and van der Does 1996 et les travaux qu'ils citent). Nous supposons que tout mécanisme qui rend compte de la distributivité d'une phrase telle que *Les étudiants de ma classe sont (tous) intelligents* devrait être capable d'expliquer la distributivité des phrases du type de (16) et (17). Comme *être intelligent* est en outre un prédicat non-spécifiant, et que la seule restriction sur s est que z est dans s , nous avons une correspondance bi-univoque entre individus et situations (cf. de Swart 1996 pour une élaboration de cette idée). On peut alors dire que dans les DRS K_1 et K_2 , l'opérateur GEN quantifie sur des situations qui ne contiennent ni plus ni moins que l'information venant du nom qui introduit le rd, ce qui aboutit au liage indirect du rd introduit par le SN sujet.

Le trait [+Max] du rd dans les deux DRS K_1 et K_2 dérive de la maximalité de l'ensemble d'événements qui constitue la restriction de l'opérateur GEN. Comme nous interprétons GEN comme un défini pluriel, l'ensemble des événements constitue la somme maximale du treillis. Le liage indirect conduit à une quantification sur l'ensemble entier des chiens. La seule différence entre les DRS K_1 et K_2 est la présence du point d'exclamation dans la DRS K_1 , qui indique la référence déterminée du rd introduit par le SN défini en français. La maximalité du rd x est donc affirmée dans K_1 . Dans K_2 , la maximalité du rd n'est pas affirmée, car nous avons affaire à un pluriel nu, qui est assimilable à l'indéfini, que nous avons caractérisé comme le membre non-marqué de la paire. Le fait que l'indéfini n'affirme rien veut dire qu'il est neutre par rapport à la référence déterminée ou non-déterminée du rd. La structure K_2 convient à l'expression de la généralité, car - grâce au liage indirect par l'opérateur GEN - elle implique la maximalité du rd x , sans toutefois la rendre explicite dans la forme du SN.

En fin de compte, les sens exprimés par K_1 et K_2 sont équivalents, même si les parties constituantes et la construction de la DRS ne sont pas identiques. La conclusion plus générale que nous pouvons tirer de cette analyse est que le contraste sémantique entre définis et indéfinis est neutralisé dans les contextes génériques, grâce au rôle de l'opérateur GEN. En somme, les SN définis, tout comme les SN indéfinis, permettent de récupérer l'input des tableaux 1 et 2 dans la section 4 ci-dessus. Il est important de noter que les traits de Maximalité et de Familiarité ne sont pas représentés en tant que tels dans la DRS. La propriété de maximalité concerne la dénotation du rd dans le modèle d'interprétation plutôt que sa représentation dans la DRS. La propriété de Familiarité est également déduite, plutôt que représentée dans la DRS. L'absence d'une condition du type $x = y$, qui identifie le rd concerné avec un autre rd déjà introduit dans la DRS, suffit pour lui accorder le trait [-Fam]. Les traits

de familiarité et de maximalité, qui constituent l'input sémantique des tableaux de la syntaxe OT de la section 2, et qui déterminent le choix de l'article dans les contextes génériques, sont donc récupérables à partir de la DRS, sans faire directement partie de la DRS.

L'argumentation s'étend aux prédicats qui expriment la référence à l'espèce. La représentation des phrases (18) et (19) en DRT donne lieu aux DRS K_3 et K_4 :

- (18) A dinoszauruszok eltűntek.
Les dinosaures ont disparu.

DRS K_3

$!x_k, y$ $Pl(y)$ $x_k = \cap \text{Dinosaure}(y)$ $\text{Disparu}(x_k)$

- (19) Dinosaurs are extinct.
Dinosaures ont disparu.

DRS K_4

x_k, y $Pl(y)$ $x_k = \cap \text{Dinosaure}(y)$ $\text{Disparu}(x_k)$
--

Dans les deux DRS, la flexion nominale plurielle introduit la présupposition qu'il existe un rd pluriel. Ce rd présupposé est lié par le rd y introduit par l'article défini dans la DRS K_3 ; dans le cas du pluriel nu, il est accommodé dans la DRS K_4 . Comme le prédicat porte sur l'espèce, il faut construire l'espèce à partir du SN pluriel défini ou indéfini. Etant donné que l'espèce est conçue comme le sommet du treillis (voir section 3 ci-dessus), nous utilisons l'opérateur Down \cap de Chierchia (1998) pour créer un individu (l'espèce) à partir d'une propriété (donnée par le contenu descriptif du nom). Dans les DRS K_3 et K_4 , \cap est utilisé pour définir x_k . Evidemment, les réalisations de l'espèce dans ce cas-ci sont des individus existant dans des mondes différents, et le treillis qui supporte x_k constitue une classe ouverte. A moins que x_k n'ait été mentionné dans le discours précédent (voir section 6.2 ci-dessous), x_k porte le trait [-Fam], car les réalisations de l'espèce sont accessibles, mais ne sont pas connues dans le contexte.⁷

La seule différence entre les deux DRS réside dans le fait que le rd x_k de K_3 est marqué comme ayant une référence déterminée ($!x_k$), tandis que le rd x_k dans K_4 ne porte aucune marque. Cette différence n'entraîne aucune conséquence sémantique en ce qui concerne les conditions de vérité de la DRS. Nous savons que x_k est unique, car chaque treillis a une seule somme maximale. Que cette valeur unique du rd soit affirmée (par le SN défini) ou pas (par le SN indéfini) ne change rien au modèle

⁷ Nous supposons que l'accommodation du rd est bloquée dans les contextes génériques. La question des restrictions sur le processus d'accommodation ne sera pas approfondie ici.

vérifiant la phrase. Nous nous retrouvons donc une fois de plus dans une situation où le contraste entre SN défini et indéfini est neutralisé par le contexte générique. Les deux types de SN peuvent servir à l'expression de la référence à l'espèce, le choix de l'article dans les contextes génériques est donc sujet à des considérations de nature syntaxique, plutôt que sémantiques.

En résumé, la généralisation générique et la référence à l'espèce constituent deux contextes où le contraste entre SN définis et indéfinis est neutralisé par les propriétés de la construction, qui entraîne la nature maximale du rd, qu'elle soit affirmée par le SN ou pas. Comme les traits sémantiques du rd sont identiques dans les deux cas, et que cette configuration implique un conflit entre les deux propriétés essentielles des SN définis, la syntaxe OT peut déterminer le choix de l'article par l'ordre hiérarchique entre les contraintes FMax et *Def/-Fam sans risque de compromettre le sens, qui est récupérable indépendamment du choix de l'article. La différence entre l'anglais et les langues germaniques d'une part et les langues romanes, le grec et le hongrois de l'autre, se réduit donc à une différence syntaxique qui n'entraîne pas de conséquences pour la sémantique.

Notre analyse prédit l'inexistence d'une langue ayant un 'article générique'. Nous avons montré que le contraste entre article défini et indéfini est sémantiquement neutralisé dans les contextes génériques, n'importe quelle forme d'article pouvant faire l'affaire: on n'a donc pas besoin d'un article supplémentaire.

6. La généricité non-standard

Jusqu'à présent, cet article a traité de deux formes 'standard' de généricité, à savoir la généralisation générique et la référence à l'espèce. Nous avons soutenu l'hypothèse que la configuration de traits [+Max] et [-Fam] du rd générique crée un conflit qui peut être résolu par le choix de l'article défini ou de l'article indéfini, selon la grammaire OT de la langue en question. Comme le sens de la phrase générique peut être récupéré dans les deux cas, le choix est essentiellement libre, et la variation qu'on trouve à travers les langues prouve que cette liberté est bien exploitée. Dans cette section, nous étendons l'analyse à deux contextes génériques non-standard qui nous permettent de tester notre hypothèse. Nous avons caractérisé l'anglais et les langues germaniques comme des langues à 'haute familiarité', car elles donnent la priorité au blocage de l'article défini pour les rd non-familiers. Par contre, les langues romanes et le hongrois sont des langues à 'haute maximalité', car elles demandent que la nature maximale du rd soit reflétée dans la forme. Cette analyse fait deux prédictions: dans un contexte générique où le rd n'a pas le trait [+Max], les langues romanes devraient choisir un article indéfini tout comme les langues germaniques. Dans un contexte générique où le rd a le trait [+Fam], les langues germaniques devraient choisir un article défini, tout comme les langues romanes. Dans les sections 6.1 et 6.2 nous traitons de deux types d'exemples qui confirment ces prédictions, à savoir la 'pseudo-généricité' (article indéfini dans toutes les langues) et la 'généricité anaphorique' (article défini dans toutes les langues).

6.1 La pseudo-généricité

Les 'vraies' généralisations génériques impliquent le liage indirect de la variable individuelle par l'opérateur générique. Nous avons utilisé les phrases du type *Les chiens sont intelligents* pour illustrer la correspondance bi-univoque entre situations et individus qui résulte de la combinaison d'un prédicat non-spécifiant et d'une restriction minimale sur les situations qui constituent le domaine de la quantification (à savoir 'x dans s'). Mais supposons maintenant qu'il y ait des détails concernant la

situation qui dépassent la simple existence de l'individu dans la situation *s*. Dans ce cas, nous perdons la correspondance bi-univoque entre individus et situations. La quantification porterait toujours sur l'ensemble des situations, mais pas nécessairement sur l'ensemble des individus. Le rd générique perdrait alors son trait [+Max], même si l'on avait toujours affaire à une phrase générique. Dans les langues qui adoptent la hiérarchie FMax >> *Def/-Fam, la configuration [-Max] [-Fam] ne peut plus entraîner l'emploi d'un article défini, car il n'y a pas de maximalité dans l'input (le sens) à refléter dans l'output (la forme). La contrainte plus faible *Def/-Fam entre ainsi en jeu, bloquant l'emploi de l'article défini. Dans ce contexte, l'article indéfini devrait par conséquent émerger comme forme optimale dans les deux classes de langues. Longobardi (2001), Dobrovie-Sorin et Laca (2003) et Dobrovie-Sorin (2004) discutent d'exemples italiens et français qui illustrent ce que nous avons défini ici comme la pseudo-généricité:

- (20) Cani da guardia di grosse dimensioni sono più efficienti.
chiens de garde de grande dimension sont plus efficaces
'Des chiens de garde de grande dimension sont plus efficaces.'
- (21) Ucelli di zone paludose sono ghiotti di insetti.
oiseaux de régions marécageuses sont avides d'insectes
'Des oiseaux de régions marécageuses sont avides d'insectes.'
- (22) Elefanti di colore bianco possono creare una grande curiosità.
éléphants de couleur blanche peuvent créer beaucoup de curiosité
'Des éléphants de couleur blanche peuvent créer beaucoup de curiosité.'
- (23) Des enfants malades sont grincheux.
- (24) Des guêpes énervées sont un danger terrible.
- (25) Des grévistes peuvent ruiner une entreprise.
- (26) Des jeunes filles doivent se montrer discrètes.

Nous pouvons donner des exemples comparables pour le hongrois:

- (27) Kibáló gyerekek fárasztóak.
Hurlant enfant.Pl fatigant.Pl
'Des enfants hurlant sont fatigans.'
- (28) Sztrájkolók tönkretelhetnek egy vállalatot.
gréviste.Pl ruiner.Poss.Pl une entreprise.
'Des grévistes peuvent ruiner une entreprise.'

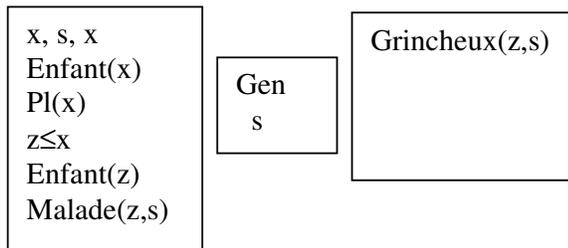
La version anglaise de ces exemples contient un pluriel nu:

- (29) a. Sick children are grouchy. [= (23), (27)]
b. Strikers can ruin an enterprise. [= (25), (28)]

Mais à la différence des exemples italiens, français et hongrois que nous venons de donner, les phrases anglaises sont ambiguës, et (29a) peut servir comme la traduction de (23) autant que de (31) ci-dessous.

Pour illustrer, nous donnons l'interprétation de (23/27/29a) en DRT :

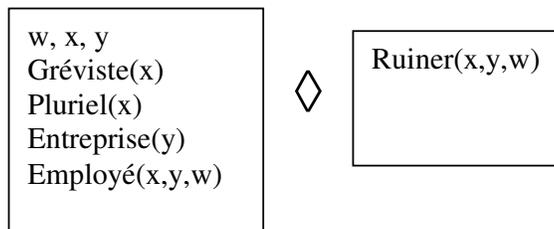
DRS K_5



L'opérateur générique lie la variable situationnelle, comme d'habitude. Tout comme dans le cas de la phrase *Les chiens sont intelligents*, nous supposons que le prédicat est de nature distributive, et porte sur les enfants individuels, plutôt que sur le groupe pluriel. Notons toutefois que le domaine de quantification dans la DRS K_5 n'est pas défini par une restriction minimale du type 'z dans s', comme on l'a vu dans les exemples illustrant le liage indirect. Nous avons maintenant affaire à une restriction plus complexe formulée comme *Malade(z,s)*. Il s'ensuit que l'opérateur générique quantifie sur les situations de maladie d'enfant. Cet ensemble de situations n'est pas dans une relation de correspondance bi-univoque avec l'ensemble des enfants. Le rd x ne porte donc pas le trait [+Max]. Dans la grammaire OT, FMax n'est pas violé quand il n'y a pas de maximalité du rd (sens) à refléter dans la syntaxe (forme). Dans ce cas entre en jeu la contrainte plus faible *Def/-Fam, qui bloque l'emploi de l'article défini pour les rd non-familiers. Etant donné que le rd x porte le trait [-Fam] comme tous les rd génériques qu'on a considérés jusqu'ici, l'article défini est bloqué par cette contrainte, et la grammaire OT choisit l'article indéfini comme la forme optimale dans le contexte de (23/27). Evidemment, la traduction de cet exemple en anglais (29a) contient un pluriel nu, car la contrainte *Def/-Fam est la contrainte la plus importante dans cette langue de toute façon. Le contraste entre les deux classes de langues est donc neutralisé dans le contexte de la pseudo-généricité, exactement comme le prédit notre analyse.

Nous proposons d'analyser les exemples modaux du type de (22), (25), (28), (29b) de la même façon. La DRS K_6 donne notre interprétation de la phrase:

DRS K_6



Grâce au verbe modal *pouvoir*, la quantification porte maintenant sur les mondes possibles (variable w). Le symbole \diamond de la modalité à valeur existentielle fonctionne donc comme l'élément qui lie les deux boîtes subordonnées. Tout comme dans la DRS K_5 le domaine de quantification n'est pas défini par une restriction minimale du type 'x dans w', comme dans les exemples illustrant le liage indirect. La subdivision arrière-plan/ focus met *ruiner* dans la portée de la quantification, tandis que tout le reste de la phrase - y compris la présupposition que seuls les grévistes qui sont employés de l'entreprise en question peuvent la ruiner - va dans la restriction. La quantification porte donc sur les mondes possibles w où des grévistes x sont employés d'une entreprise y . Dans cette configuration, il n'y a pas de correspondance un à un entre l'ensemble des grévistes et les mondes w . Bien sûr, le quantificateur porte sur l'ensemble (maximal) des mondes pertinents, mais l'ensemble des grévistes ne

constitue pas le domaine de quantification, ni directement, ni indirectement. Le rd x ne porte donc pas le trait [+Max], mais seulement le trait [-Fam]. Par conséquent, l'article défini est bloqué, en français autant qu'en anglais.

Le contraste entre généricité et pseudo-généricité est assez subtil, et souvent les deux formes sont possibles. Ainsi, la phrase (30) existe à côté de (26), mais n'exprime pas tout à fait le même sens (Carlier 1989):

(30) Les jeunes filles doivent être discrètes.

Selon Carlier, la version avec article défini constitue une généralisation descriptive, tandis que la version avec le SN indéfini (ex. (26)) a une valeur déontique. Le fait que les jeunes filles soient prises comme un ensemble global, sous-espèce de l'espèce des filles dans (30) explique pourquoi la phrase exprime une généralisation générique. La phrase (26) par contre porte sur des filles en situation de jeunesse. Cette construction du sens met l'accent sur la localisation dans le temps, et entraîne la valeur de règle de conduite.

De même, la phrase (31) existe à côté de (23):

(31) Les enfants malades sont grincheux.

Tout comme dans (30), l'article défini construit une sous-espèce des enfants, et maintient la nature maximale du rd générique. Le choix entre (24) et (31) est déterminé par la question de savoir si on veut exprimer une généralisation générique sur la sous-espèce d'enfants malades (article défini) ou mettre l'accent sur les situations de maladies d'enfants (article indéfini) (cf. Heyd 2003). Heyd fait observer que la présence d'un adjectif tel que *malade* est indispensable dans la construction de la pseudo-généricité, et que seuls les adjectifs à valeur de prédicat spécifiant (cf. Kleiber 1981), impliquant une localisation spatio-temporelle, sont capables de créer cet effet de sens. Dans notre système, cette restriction s'explique par le fait que l'absence d'une spécification de la situation mènerait à un cas de liage indirect, donc à l'emploi d'un article défini.⁸

En l'absence d'un contraste formel, la différence de sens entre la généricité standard et la 'pseudo-généricité' ne se fait pas sentir en anglais autant qu'en français ou dans les autres langues romanes. C'est que l'anglais (29a) correspond à la fois à (23) et à (31) en français. En d'autres termes, la nature maximale de l'ensemble des individus dans l'expression d'une généralisation générique n'est pas reflétée dans la forme de toute façon. L'ambiguïté potentielle des pluriels nus de l'anglais dans les contextes générique a été signalée par Carlson et Pelletier (1995), et les observations faites par Longobardi (2001) vont dans la même direction. La discussion lancée par Chierchia (1998), Longobardi (2001), Dobrovie-Sorin et Laca (2003) et Dobrovie-Sorin (2004) a mis en valeur les cas de pseudo-généricité pour les langues romanes. L'existence d'exemples du même type en hongrois montre bien qu'il ne s'agit pas d'une particularité des langues romanes, mais que le choix de l'article est sensible aux traits de Maximalité et de Familiarité du rd générique. La neutralisation du contraste entre l'anglais et les langues germaniques, d'une part, et les langues romanes et le hongrois, de l'autre, étaye notre analyse dans le cadre de la théorie de l'optimalité.

⁸ Nous avons l'impression que dans le cas de la pseudo-généricité, les nuances de sens varient légèrement d'une langue à l'autre, mais nous nous limitons ici à la description la plus générale possible.

6.2 La généricité anaphorique

Dans les exemples de pseudo-généricité que nous venons de discuter, le contraste entre l'anglais et les langues germaniques, d'une part, et les langues romanes et le hongrois, de l'autre, disparaît à cause du trait [-Max] du rd générique, qui bloque l'emploi de l'article défini même dans les grammaires qui adoptent la hiérarchie FMax >> *Fam/-Def. Dans la configuration de traits donnée par la pseudo-généricité, on voit ainsi l'article indéfini émerger dans toutes les langues comme la forme préférée. Dans cette section, nous traitons d'un deuxième cas de neutralisation des formes, cette fois-ci en faveur de l'article défini dans toutes les langues. Nous parlons de généricité anaphorique pour décrire les cas où le rd générique porte les traits [+Max] et [+Fam]. On peut trouver des exemples de ce type de généricité dans les entrées encyclopédiques de noms tels que *dinosaure*. Une fois introduite l'espèce des dinosaures, le rd générique a un antécédent dans le discours, et porte le trait [+Fam]. Dans ce cas, on trouve l'article défini, comme l'illustrent les exemples suivants de l'anglais:

- (32) *Saurischian Bipedes* – *The Saurischians* were the first of the two great groups to assume prominence. [...] From certain of these forms, *the Saurischians* were certainly derived. (Encyclopedia Britannica, 1972, p. 456).
'Bipèdes sauriens — Les Sauriens furent le premier des deux grands groupes à prendre une certaine importance. [...] Les Sauriens se sont certainement développés à partir de certaines de ces formes.'
- (33) The Mesozoic, the geologic scene in which *the dinosaurs* played the most important part, extended ... (from Encyclopaedia Britannica, 1972, p. 454)
'Le mésozoïque, l'ère géologique pendant laquelle les dinosaures jouèrent le rôle le plus important, s'étendit...'
- (34) *The dinosaurs* belong to a major subdivision of the reptiles termed the Archosauria, or ruling reptiles. (From Encyclopaedia Britannica, 1972, p. 455)
'Les dinosaures appartiennent à l'une des branches principales de la famille des reptiles, celle des archosaures, ou reptiles dominants.'

Les exemples de ce type contiennent souvent des prédicats exprimant la référence à l'espèce, comme *dérivée*, *appartenir à une branche de*, etc. Mais (35) indique que nous pouvons trouver également des lectures habituelles de prédicats non-spécifiants:

- (35) *The coelurosaurs* ... were slender-limbed, lightly-built ... and obviously fast-running forms that appear to have preyed upon their smaller reptilian contemporaries.
'Les coelurosaures...étaient des animaux à membres fins et à squelette léger..., manifestement rapides à la course, qui semblent avoir chassé les reptiles de taille inférieure.'

Ces exemples diffèrent de la généricité standard en ce que l'espèce ou la somme des individus réalisant l'espèce a déjà été introduite dans le contexte précédent. Le rd générique a donc un antécédent dans le contexte discursif, ce qui en fait un rd qui porte le trait [+Fam]. À part sa nature anaphorique, ce rd fait clairement référence à l'espèce, et porte donc le trait [+Max]. Dans une grammaire où la hiérarchie *Def/-Fam >> FMax bloque l'article défini pour les cas de généricité standard, la valeur [+Fam] du rd générique fait toute la différence. La contrainte la plus importante n'est plus violée par l'input, et la contrainte FMax est maintenant

décisive. Comme le rd générique porte le trait [+Fam], même l'anglais et les langues germaniques optent pour l'article défini dans les cas de généricité anaphorique. L'emploi fréquent de l'article défini dans les entrées encyclopédiques dans toutes les langues est donc naturellement expliqué par notre grammaire OT.⁹

7. Conclusion

L'analyse que nous avons présentée rend compte du choix de l'article dans les contextes génériques dans deux classes de langues, illustrées par l'anglais et les langues germaniques d'une part, et les langues romanes, le grec et le hongrois, de l'autre. Nous avons formulé deux contraintes universelles, qui sont fondées sur la sémantique des SN définis et indéfinis proposée par Farkas (2002), à savoir FMax ('réfléter la maximalité du référent discursif dans la forme') et *Def/-Fam ('éviter les définis non-familiers'). Les deux ordres hiérarchiques possibles entre les deux contraintes dans une grammaire OT rend compte des deux classes de langues. La hiérarchie FMax >> *Def/-Fam qu'on trouve dans les langues romanes et en hongrois donne la priorité à l'article défini dans les contextes de référence à l'espèce et de généralisation générique. La hiérarchie *Def/-Fam >> FMax bloque l'emploi de l'article défini dans ces contextes en anglais et dans les langues germaniques en général. Notre analyse prédit que le contraste entre les deux classes de langues est neutralisé lorsque la configuration de traits est modifiée. La pseudo-généricité illustre le fait que toutes les langues optent pour l'article indéfini comme forme optimale quand le rd générique ne porte pas le trait [+Max]. La généricité anaphorique montre que l'article défini est de mise dans toutes les langues pour les cas où le rd générique porte le trait [+Fam]. Ces résultats confirment l'approche générale que nous défendons dans cet article, à savoir que le choix de l'article dans les contextes génériques relève d'une paramétrisation syntaxique (qui est déterminée par la grammaire OT), mais n'affecte pas la sémantique, qui est identique dans le cas des SN génériques définis et indéfinis.

⁹ Dans ce volume, Kwon et Zribi-Hertz traitent d'un autre exemple de généricité définie dans les langues germaniques, qui va de pair avec un emploi particulier du marqueur de pluralité *deul* en coréen. Il s'agit du contraste entre le pluriel indéfini de phrases telles que (i) et le pluriel défini de phrases telles que (ii):

- (i) Germans like music.
- (ii) The Germans like music.
- (iii) Les Allemands aiment la musique.

(i) est une vraie phrase générique, comportant un pluriel nu. Le SN défini de (ii) par contre, renvoie à l'ensemble fermé des gens qui constituent la nation allemande, quel que soit le moment, et sans qu'on ait accès aux individus constituant l'ensemble. Les deux interprétations se traduisent par (iii) en français. Dans notre approche, (ii) implique une interprétation spéciale du substantif, plutôt que de l'article. Dans ce type d'exemple, où le substantif se prête à une interprétation fonctionnelle ou intensionnelle, la phrase peut prendre une valeur généralisante sans qu'il s'agisse d'une vraie phrase générique. Nous laissons les phrases de ce type donc de côté dans notre article, et nous renvoyons le lecteur à Kwon et Zribi-Hertz (ce volume) pour une analyse du contraste entre la généralisation nominique exprimée par (i), et la généralisation accidentelle exprimée par (ii).

Bibliographie

- Beaver, D. (2004). The optimization of discourse anaphora *Linguistics and Philosophy* 27, 3-56.
- Beyssade, C. (2005). Les définis génériques en français: noms d'espèces ou sommes maximales. (ce volume)
- Bosveld-de Smet, L. (1998). On mass and plural quantification: the case of French *des/du* NPs, PhD thesis, University of Groningen.
- Carlier, A. (1989). Généricité du syntagme nominal sujet et modalités, *Travaux de Linguistique* 19, 33-56.
- Carlson, G. (1980). *Reference to kinds in English*, New York: Garland.
- Carlson, G. and J. Pelletier (1995) (eds.). *The generic book*, University of Chicago Press.
- Chierchia, G. (1998). Reference to kinds across languages, *Natural Language Semantics* 6, 339-405.
- Corblin, F. (2002). *Représentation du discours et sémantique formelle. Introduction et applications au français*, Paris: PUF.
- Dayal, V. (2004). Number marking and (in)definiteness in kind terms, *Linguistics and Philosophy* 27, 393-450.
- Dobrovie-Sorin, C. (1997). Classes de prédicats, distribution des indéfinis, et la distinction thétiq-ue-catégorique, *Le gré des langues*.
- Dobrovie-Sorin, C. et B. Laca (1999). La généricité entre la référence à l'espèce et la quantification générique, *Actes de Langue et Grammaire* 3, Presses Universitaires de Vincennes.
- Dobrovie-Sorin, C. et B. Laca (2003). Les noms sans déterminant dans les langues romanes, in: D. Godard (éd.) *Les langues romanes-Problèmes de la phrase simple*. Eds. du CNRS, 235-281.
- Dobrovie-Sorin, C. (2004). Generic plural indefinites and (in)direct binding, in: F. Corblin et H. de Swart (éds.) *Handbook of French Semantics*, Stanford : CSLI Publications.
- Farkas, D. (2002). Specificity distinctions, *Journal of Semantics* 19: 1-31.
- Farkas, D. et H. de Swart (2003). *The semantics of incorporation : from argument structure to discourse transparency*. Stanford: CSLI Publications.
- Farkas, D. et H. de Swart (2005). The generic article, in: H. de Hoop et J. Zwarts (éds.). *Proceedings of the workshop on cross-linguistic semantics*, Edinburgh: ESSLLI summerschool.
- Ferreira, M. (2005). Bare habituals and plural definite descriptions, *Proceedings of Sinn und Bedeutung* 9, Nijmegen, 102-115. See: <http://www.ru.nl/ncs/sub9/>
- Hawkins, J. (1991). On (in)definite articles: implicatures and (un)grammaticality predictions, *Journal of Linguistics* 27, 405-442.
- Heyd, S. (2003). *L'interprétation des syntagmes nominaux en "des" et "de" en position sujet et objet - Généricité, habitualité et incorporation sémantique*, PhD thesis, Université Marc Bloch (Strasbourg).
- Kamp, H. et U. Reyle (1993). *From discourse to logic*. Dordrecht : Kluwer Academic Press.
- Kleiber, G. (1981). Relatives spécifiantes et non-spécifiantes, *Le français moderne* 49, 216-233.
- Kwon, S.-N. et A. Zribi-Hertz (2005). Pluriel et généricité nominale: les enseignements du coréen (ce volume).

- Lewis, D. (1979). Score keeping in a language game, in: R. Bäuerle et al. (eds.). *Semantics from different points of view*, Berlin : Springer, 172-187.
- Link, G. (1983). The logical analysis of plurals and mass terms : a lattice-theoretic approach. In: R. Bäuerle et al. (eds.). *Meaning, use and interpretation of language*, Berlin : de Gruyter, 302-323.
- Longobardi, G. (2001). How comparative is semantics? A unified parametric theory of bare nouns and proper names, *Natural Language Semantics* 9 : 335-369.
- Longobardi, G. (1994). Reference and proper names, *Linguistic Inquiry* 25 : 609-665.
- Ojeda, A. (1993). *Linguistic individuals*, Stanford : CSLI Publications.
- Robertson, H. (2004). Definites, deixis and disclosure : the semantics of the French plural generic definite, talk presented at GURT 2004, Georgetown University.
- Roodenberg, J. (2004). French bare arguments are not extinct: the case of coordinated bare nouns, *Linguistic Inquiry* 35, 303-313.
- van der Sandt, R. (1992). Presupposition projection as anaphora resolution, *Journal of Semantics* 9, 333-377.
- de Swart, H. (1991). *Adverbs of quantification: a generalized quantifier approach*. PhD. Dissertation Groningen Univ.; published 1992 by Garland, New York.
- de Swart, H. (1996). (In)definites and genericity, in : M. Kanazawa et al. (eds.). *Quantifiers, deduction and context*, CSLI Publications, p. 171-194.
- Vergnaud, J.-R. et M.-L. Zubizarreta (1992). The definite determiner and the inalienable construction in French and English, *Linguistic Inquiry* 23: 595-652.
- Verkuyl, H. et J. van der Does (1996). The semantics of plural noun phrases, in : J. van der Does et J. van Eijck (eds.). *Quantifiers, logic and language*, Stanford : CSLI Publications, 337-374.