

La description de la musique

Cet article n'a pas pour ambition de dresser un panorama exhaustif des différentes manières de décrire le contenu des documents musicaux dans les bibliothèques et centres de documentation, les pratiques locales pouvant varier à l'infini.

Isabelle Gauchet

Centre de documentation de la musique contemporaine (Cdmc)
isabelle.gauchet@cdmc.asso.fr

Son objet est de montrer en quoi les structures proposées par les formats normalisés à ce jour ne sont pas satisfaisantes et de présenter un exemple de solution mise en œuvre par trois organismes, qui concerne essentiellement la description instrumentale et vocale. L'analyse de leurs pratiques conduira ensuite à envisager quelques perspectives nouvelles, liées en particulier à l'évolution du format international Unimarc.

Les principaux formats normalisés

Rappelons, tout d'abord, quelles sont les principales possibilités de description de la musique offertes par deux formats normalisés utilisés couramment dans les bibliothèques et centres de documentation en France : Unimarc et Interarc.

Le format international Unimarc propose la zone de données codées 125 pour décrire, notamment, la nature de la musique imprimée¹ ; cette information, qui concerne plutôt la forme du document, peut donner en même temps une information sur son contenu et la nature de la musique elle-même, puisque certains codes permettent d'indiquer s'il s'agit d'une partition d'orchestre, d'une partition de chant avec accompagnement réduit pour le piano, d'une partition

vocale ou chorale sans accompagnement...

La zone 500 du titre uniforme possède, quant à elle, quelques sous-zones spécifiques à la description de la musique : la sous-zone \$r où est indiqué le mode d'interprétation, c'est-à-dire l'instrumentation ; la tonalité est indiquée dans la sous-zone \$u ; la sous-zone \$w permet d'indiquer s'il s'agit d'un arrangement.

Le bloc 6XX de l'indexation matière offre de grandes libertés d'indexation pour la musique, notamment par l'utilisation de la zone 606, nom commun - vedette matière, qui doit cependant respecter un système d'indexation donné, que ce soit un thésaurus ou une liste d'autorité, Rameau (Répertoire d'autorité-matière encyclopédique et alphabétique unifié) par exemple.

La seule zone prévue spécifiquement par Unimarc pour décrire le contenu de la musique est, en fait, la zone de données codées 128, interprétations musicales et partitions : elle décrit la forme de la composition musicale et l'instrumentation et/ou la vocalisation. Elle comprend trois sous-zones : la première sous-zone (128 \$a), répétable, reçoit sous la forme d'un code de deux caractères la forme de la composition, *co* pour concerto ou *or* pour oratorio par exemple. La liste comprend actuellement quelque soixante-dix codes.

Les deux autres sous-zones (\$b et \$c), répétables, qui servent à décrire les instruments et les voix, se construisent de la même façon et uti-

1. Cette zone permet aussi de décrire sous forme codée le type de texte écrit des enregistrements parlés.

Documentaliste, responsable du service documentation du Centre de documentation de la musique contemporaine, **Isabelle Gauchet** est titulaire d'un DEA Sciences de l'Antiquité et du diplôme supérieur de l'INTD. Elle a coécrit des articles sur la description instrumentale et vocale dans les revues Documentaliste-Sciences de l'information et Écouter Voir.

lisent les mêmes codes, chacune étant codée sur deux caractères, suivis éventuellement de deux caractères numériques pour préciser le nombre d'instruments ou de voix : la sous-zone \$b s'applique aux instruments ou voix dans un ensemble, la sous-zone \$c aux instruments ou voix pour solistes. Une liste de quatre-vingt-huit codes est spécifiée pour ces deux sous-zones, répartis entre dix catégories. Une pièce pour deux guitares sera ainsi codée en \$b *tb02* ; un concerto pour flûte et orchestre de chambre : \$b *ob* \$c *wa01*.

Le nouveau format InterMarc de la BnF, s'il diffère du précédent de par les étiquettes de ses zones (144 pour le titre uniforme musical, par exemple), offre cependant des possibilités voisines, notamment si l'on considère la zone de données codées 048 : cette zone, prévue pour décrire l'effectif vocal et instrumental de l'œuvre, comprend deux sous-zones répétables, \$a pour un exécutant ou un ensemble, \$b pour un soliste ; elles sont de longueur fixe (4 positions), les positions 0 et 1 recevant un code à deux caractères d'instrument ou de voix, les positions 2 et 3 recevant un nombre de deux caractères. Ainsi une sonate pour violon et piano se verra codée \$a *sa01* \$a *ka01*.

Les zones de données codées existantes posent divers problèmes et sont finalement peu utilisées dans la pratique. Les listes de codes sont le reflet des travaux effectués jusqu'à une période récente (début du XX^e siècle) sur la tradition musicale savante. Outre leur caractère relativement sibyllin au premier abord, elles ne sont donc pas suffisamment précises et étendues, en particulier lorsqu'il s'agit de décrire des musiques autres que la musique classique occidentale

(musique contemporaine, musique traditionnelle, musique extra-européenne...). Ainsi il existe un code unique pour toutes les guitares (tb), et aucun pour l'accordéon. Par ailleurs, aucune sous-zone n'est prévue pour recevoir les nombres d'instruments et de voix, qu'il s'agisse des nombres par famille ou des nombres totaux.

Une pratique actuelle : les champs 9XX

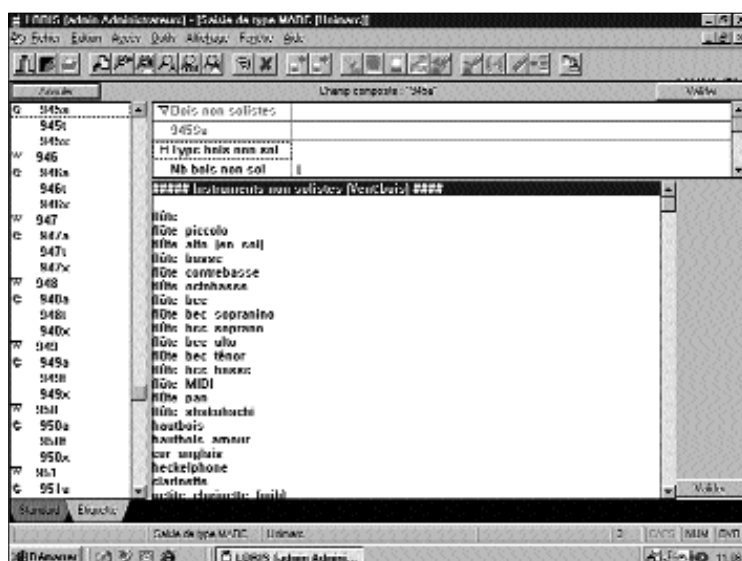
Trois organismes musicaux, le Centre de documentation de la musique contemporaine (Cdmc), la Médiathèque pédagogique de la Cité de la musique et la Médiathèque de l'Ircam (Institut de recherche et coordination acoustique/musique) ont souhaité convertir leur catalogue au format Unimarc à peu près au même moment. Leur objectif était bien sûr d'uniformiser les méthodes de description instrumentale des documents musicaux (partitions, enregistrements sonores, etc.), en vue de permettre l'échange de notices.

Mais il s'agissait aussi d'offrir à l'utilisateur des possibilités de recherches performantes sur l'instrumentation, avec des réponses claires, lisibles, et immédiatement compréhensibles. Les trois centres, rencontrant les mêmes difficultés pour trouver un système de description de leurs documents musicaux suffisamment précis et performant, ont constitué un groupe de travail et ont élaboré, à partir d'un logiciel documentaire au format Unimarc², une solution originale. La logique du système Unimarc a été prise comme point de départ ; elle consiste à détailler et structurer au maximum l'information dans des champs et sous-champs spécifiques. C'est le bloc 9XX, prévu par Unimarc pour les « données locales » et les besoins spécifiques de chaque bibliothèque, qui abrite la description de la nomenclature musicale.

2. Il s'agit du logiciel Loris de la société Ever.

Quatorze champs, de 940 à 953, ont été définis pour recevoir les voix solistes, les chœurs, les instruments solistes, les différentes familles d'instruments non solistes, d'autres informations relatives à la gestique (danseurs, acteurs...) ou à la direction. Chacun d'eux contient trois sous-champs, qui reçoivent respectivement les voix ou instruments requis, le nombre de voix/d'instruments pour le champ en question, et une note destinée à fournir si nécessaire des précisions supplémentaires. Le champ 954 a été structuré en trois sous-champs pour les effectifs totaux de voix, d'instruments non solistes et de tous les instruments. Le champ 956, conçu sur le même modèle que les champs 940 à 953, et dédié au type d'ensemble ou d'orchestre, permet une description à un niveau plus global. Pour les zones 940 à 953 et 956, la saisie des sous-champs \$a est assistée grâce à des listes établies par familles d'instruments ou de voix et adaptées à chaque zone. Chaque sous-champ \$a est répétable et possède une structure composite qui permet la concaténation du type d'instrument (ou voix) et de son nombre. La sélection de ces sous-champs composites donne accès à une fenêtre, sous laquelle apparaît la liste des instruments (ou voix) correspondant au champ en question, comme le montre la capture d'écran ci-jointe.

Après sélection et validation d'un terme de la liste, il suffit de préciser le nombre (la valeur 1 étant la valeur par défaut). Dans le sous-champ suivant (\$t), est saisi le nombre de voix ou d'instruments de la famille décrite. Une fois toute l'instrumentation saisie, les nombres totaux de voix et d'instruments pour l'œuvre contenue dans le document catalogué sont indiqués dans le champ 954. L'assistance sur liste permet d'éviter les erreurs lors de la saisie. Ces listes, établies en commun par les trois centres afin de permettre éventuellement l'échange de notices, doivent demeurer absolument identiques et



Capture d'écran. Exemple de sous-champ \$a à structure composite.

ne peuvent donc être modifiées sans l'accord de chacun d'eux. Un champ « candidat instrument » a été en outre prévu pour pouvoir saisir un instrument non encore disponible sur ces listes. Cependant, dans la pratique, il est peu utilisé, l'enrichissement coordonné des listes entre les trois organismes se faisant avec beaucoup de souplesse et de rapidité.

Le tableau ci-contre détaille la structure des champs utilisés pour la description instrumentale et vocale, ainsi que deux champs complémentaires définis au sein du même bloc 9XX pour les informations concernant la composition et la création de l'œuvre.

Tous les sous-champs de la description instrumentale et vocale étant indexables dans le système, divers types de recherche sont possibles, de la recherche monocritère (œuvres avec flûte par exemple) à des recherches multicritères complexes combinant les opérateurs booléens « et », « ou », « sauf ». On peut ainsi chercher les œuvres pour soprano *ou* haute-contre avec cinq à dix instruments ; les œuvres pour chœur de femmes *et* orchestre ; les œuvres pour cinq instruments *sauf* la harpe... avec un taux de pertinence des réponses excellent.

Les recherches sur la description instrumentale et vocale sont également possibles *via* l'Internet grâce à une interface³, l'objectif des trois centres étant de trouver un compromis entre une palette de critères suffisamment large et une utilisation relativement simple pour tout utilisateur.

Évolutions possibles d'Unimarc

Des perspectives d'évolution importante dans la description de la musique sont actuellement à l'ordre du jour. Une double réflexion s'est en effet engagée au sein de deux groupes de travail respectivement français et italien : le groupe « documents sonores » mis en place par le Comité français Unimarc (*cf.* l'encadré de Jean-Paul Gaschignard), et le sous-groupe Musique du comité italien Unimarc. Les comités nationaux soumettent ensuite leurs propositions

3. Pour la médiathèque pédagogique de la Cité de la musique : <http://mediatheque.cite-musique.fr/> ; pour la médiathèque de l'Ircam : <http://mediatheque.ircam.fr/> ; pour le Cdmc : <http://www.cdmc.asso.fr> rubrique « Le Catalogue ».

au *Permanent Unimarc Committee* (PUC) de l'Ifla dans le but de faire évoluer le format Unimarc pour mieux répondre aux besoins.

Plusieurs des propositions italiennes concernant la description de la musique sont étudiées avec attention par le CFU (Comité français Unimarc). Parmi celles-ci, citons brièvement la création d'une zone 136 pour l'*incipit* musical, qui recevrait sous forme codée les premières mesures de l'œuvre musicale et, si nécessaire, les premiers mots du texte mis en musique ; la création d'une zone 621, Informations sur l'exécution, qui recevrait les informations concernant l'exécution de l'œuvre (lieu, date, ...).

La description du contenu musical est le point le plus complexe, sur lequel le groupe français porte une attention particulière. Il a fait l'objet d'une étude approfondie de la part du groupe italien Unimarc, qui a relevé lui aussi le manque de précision et les lacunes de la norme existante. Après avoir étudié avec intérêt le système des zones 9XX mis en place par le Cdmc, l'Ircam et la Cité de la musique, il a élaboré une première proposition destinée à pallier les lacunes de la zone de données codées 128 : réserver celle-ci à la description de la forme (ou du genre) de la composition avec suppression des sous-zones \$b et \$c ; créer une nouvelle zone de données codées pour décrire la distribution d'exécution, c'est-à-dire le contenu instrumental et vocal des documents musicaux. Le groupe français a émis un avis favorable sur le principe : il lui semblait effectivement utile de séparer ces deux types de données et de mieux les structurer pour qu'elles puissent être utilisées plus efficacement.

Le comité français avait également travaillé de son côté sur l'enrichissement des codes destinés à décrire la forme ou le genre de l'œuvre musicale en 128 \$a, afin de créer des codes pour certaines formes spécifiques, et d'autres correspondant aux nouveaux types de musique. Cependant, la proposition italienne de créa-

Tableau. Description instrumentale et vocale
Structure des champs

Étiquettes et intitulés des champs	S	Description des sous-champs
940 Voix solistes	\$a \$t \$x	type de voix et nombre nombre total de voix solistes notes
941 Chœurs	\$a \$b \$t \$x	type de chœurs et nombre formation du chœur nombre total de voix du chœur notes
942 Instruments solistes	\$a \$t \$x	type d'instrument soliste et nombre nombre total d'instruments solistes notes
943 Gestique	\$a \$t \$x	type de gestique et nombre nombre total de gestique notes
944 Direction	\$a \$t \$x	type de chef et nombre nombre total de chefs notes
945 Instruments-bois	\$a \$t \$x	type de bois et nombre nombre total de bois notes
946 Instruments-saxophones	\$a \$t \$x	type de saxophones et nombre nombre total de saxophones notes
947 Instruments-cuivres	\$a \$t \$x	type de cuivres et nombre nombre total de cuivres notes
948 Instruments-percussions	\$a \$t \$x	type de percussions et nombre nombre total de percussions notes
949 Instruments-claviers	\$a \$t \$x	type de claviers et nombre nombre total de claviers notes
950 Instruments-cordes pincées	\$a \$t \$x	type de cordes pincées et nombre nombre total de cordes pincées notes
951 Instruments-cordes frottées	\$a \$t \$x	type de cordes frottées et nombre nombre total de cordes frottées notes
952 Instruments-divers	\$a \$t \$x	type d'instrument divers et nombre nombre total d'instruments divers notes
953 Électroacoustique	\$a \$t \$x	type d'électroacoustique et nombre nombre total notes
954 Nombre total	\$r \$s \$t	de voix d'instruments non solistes de tous les instruments
955 Composition	\$a \$b \$c \$d \$e	date lieu durée commanditaire dédicataire
956 Ensembles	\$a \$t \$x	type d'ensemble (et/ou orchestre) et nombre nombre total d'ensembles et d'orchestres notes
957 Candidats descripteurs	\$a	nouvel instrument
980 Création de l'œuvre	\$f \$i \$j \$x	date de la création interprètes de la création lieu, événement de la création notes

tion d'une nouvelle zone de données codées posait divers problèmes et n'était pas acceptable en l'état : un simple problème de nomenclature tout d'abord, la zone proposée initialement à la création étant la zone 129. Or, toute zone contenant un 9 est réservée à une utilisation nationale ou locale et n'est donc pas définie par le Comité Unimarc permanent.

D'autres problèmes ont également été soulignés par le groupe français : le codage des voix et des instruments sur seulement deux caractères, qui s'avère très vite insuffisant pour couvrir la diversité des termes ; l'absence de sous-zones numériques pour certaines catégories d'instruments ou pour le nombre total de voix et d'instruments. Des échanges fructueux ont eu lieu à ce sujet lors du 19^e congrès de l'Association internationale des bibliothèques, archives et centres de documentation musicaux (AIBM) à Périgueux, en juillet 2001, entre membres des différents groupes⁴.

Le groupe italien a, depuis, élaboré une nouvelle proposition, qu'il a récemment présentée au PUC (décembre 2001). Cette proposition apporte des évolutions importantes par rapport à la précédente. Tout d'abord, le problème de nomenclature a été réglé : la zone de données codées proposée à la création et intitulée dans la proposition originale « Medium of performance » est la zone 145. Elle est prévue pour décrire le contenu instrumental et vocal, aussi bien des partitions imprimées et manuscrites que des enregistrements audio et vidéo.

Les codes sont désormais sur trois caractères et dérivés des termes eux-mêmes, ce qui, outre l'accroissement important du nombre de combinaisons possibles, leur confère aussi une plus grande lisibilité et un caractère mnémotechnique (*flu* au lieu de *wa* par exemple pour la flûte, ou *obo* au lieu de *wb* pour le hautbois). Un système de suffixe permet en outre de

4. Cf. le compte rendu paru dans le *BBF*, 2001, t. 46, n° 6, p. 111-114 (Ndlr).

Le groupe « documents sonores » du Comité français Unimarc

Sur proposition de la Fulbi (Fédération des utilisateurs de logiciels pour bibliothèques), le Comité français Unimarc a mis en place début 2001 un groupe de travail sur les documents sonores.

Ce groupe prépare des propositions d'évolution du format, qui seront ensuite soumises au Comité et lui serviront de bases de travail. Il étudie aussi, pour le compte du Comité, des propositions émanant d'autres pays, surtout de l'Italie, où un autre groupe traite du même sujet et des manuscrits musicaux.

Il a tenu trois réunions de travail, dont une en présence de fournisseurs de logiciels et de fournisseurs de données, et souhaite travailler – ce qu'il a en partie commencé – sur les points suivants :

- le traitement des œuvres contenues ;
- les codes de fonctions des auteurs ;
- les recherches par les instrumentations des œuvres ;
- le traitement des titres uniformes musicaux.

En plus de celles contenues dans cette liste, plus d'une dizaine de propositions sont à l'étude. Le sujet prioritaire et le plus difficile est certainement le traitement des œuvres contenues : si de plus en plus de logiciels proposent la création de notices analytiques, aucun n'est aujourd'hui en mesure d'exporter des ensembles de notices et de sous-notices. Seule la Bibliothèque nationale de France le fait avec le cédérom BnF-DSAM (documents sonores, audiovisuels et multimédias), dans un format qui sera prochainement profondément modifié, avec la mise en service des fonctions d'exportation de son Système informatique.

La structure d'imbrication de notices, définie par Unimarc depuis ses origines, n'a été que très peu appliquée, et apparemment aucun fournisseur de logiciel ne souhaite l'utiliser aujourd'hui. Le traitement des œuvres contenues doit probablement proposer à la fois une ou des solutions où chaque œuvre est décrite dans une sous-notice complète, et une solution plus légère, avec un seul niveau de notice, mais en liant néanmoins les accès titres et les accès auteurs, de façon à éviter le syndrome « Carmen-Mozart »¹. Cette solution légère correspond à la structure des zones 4XX « à plat » intégrée dans Unimarc depuis sa version 3

(variante B « technique des sous-zones classiques »²), mais celle-ci doit être complétée.

Le groupe « documents sonores » souhaitait aussi travailler sur le traitement des codes-barres commerciaux : cette question a été réglée par une proposition anglaise sur la zone 017, adoptée par le PUC (*Permanent Unimarc Committee*) au printemps dernier.

Le Comité français Unimarc sera très probablement en mesure de présenter des propositions mises en forme et argumentées au PUC à la fin de 2002.

Faire évoluer Unimarc, aujourd'hui, ne peut avoir que des buts provisoires : le temps des formats Marc est compté, puisque nous verrons arriver, dans les prochaines années, des logiciels dont les données seront structurées en XML. Mais si le format change, la structure logique qui l'organise doit être maintenue et perfectionnée, et donc intégrée dans les futures DTD (définition de type de document) de XML : les performances des logiciels, leur adaptation aux besoins des établissements et de leurs lecteurs en dépendent. Aucun utilisateur ne se satisferait d'un « progrès » technique qui aboutirait à moins de fonctionnalités, moins de facilités d'utilisation, moins de productivité, bref à une régression. On peut donc penser que les adaptations préparées par le groupe « documents sonores » et le Comité français Unimarc serviront aussi après Unimarc, et c'est ce qu'il faut souhaiter.

Jean-Paul Gaschignard

Directeur de la Bibliothèque du Cher
jp.gaschignard@cg18.fr

1. Cas de figure où tous les titres de toutes les œuvres contenues dans un document sont liés à tous les auteurs : dans une anthologie d'airs d'opéra, « Carmen » peut alors être liée à l'auteur Mozart.

2. Cf. *Manuel Unimarc : format bibliographique*. Version française / [Fédération internationale des associations de bibliothécaires et des bibliothèques, Ifla Universal Bibliographic Control and International Marc Programme, Deutsche Bibliothek, Frankfurt am Main] ; traduit par Marc Chauveinc, 3^e édition, München, K.G. Saur, 1999, p. 210-214.

donner des précisions sur les instruments et les voix (tessiture, instrument électrique ou électronique...). Des sous-zones ont été ajoutées : nombre de cordes frottées, de cordes pincées, d'instruments solistes...

En résumé, la structure finale de la zone proposée repose sur deux catégories de sous-zones. La première catégorie comporte sept caractères et

concerne neuf sous-zones répétables, \$a à \$i. Les positions 0-1 reçoivent le nombre d'instruments ou de voix ; les positions 2-4 le type d'instrument ou de voix (codé sur trois caractères) ; les positions 5-6 permettent l'ajout éventuel d'un ou deux suffixes. Ces neuf sous-zones sont établies de la manière suivante : voix solo, instruments solo, voix chorales,

instruments dans un ensemble, type de chœur, type d'ensemble, groupes instrumentaux dans un ensemble plus large, chefs, autres exécutants (mimes, danseurs, etc.) et dispositifs. Les listes de codes à utiliser sont réparties en familles de voix et d'instruments à des fins pratiques uniquement.

La seconde catégorie comporte trois caractères numériques et concerne douze sous-zones non répétables, \$j à \$u. Elles permettent de préciser les nombres de voix et d'instruments par famille et le nombre total d'instruments.

La zone 145 ainsi définie, permet différents niveaux de précision dans la description de la distribution d'exécution et peut donc s'adapter aux divers besoins locaux des agences de catalogage : d'un niveau global (description des ensembles instrumentaux ou vocaux) à un niveau analytique (détail des instruments et des voix). Malgré sa relative complexité, elle constitue certainement une voie à prendre en compte pour l'échange des données de description musicale.

Les pratiques dans la description de la musique semblent promises à de nombreux bouleversements, l'évolution d'Unimarc en constituant un volet important. En attendant l'avènement des métalangages (du type XML⁵) qui devraient permettre de simplifier considérablement l'échange de données informatisées, l'évolution de la norme conditionnera la réussite des échanges de données à un niveau international. Il reste à souhaiter, si ces propositions sont avalisées par le Comité permanent Unimarc de l'Ifla, que les évolutions du format soient rapidement prises en compte par les fournisseurs de logiciels, de manière à proposer des systèmes permettant l'interrogation des données codées relatives à la description de la musique et leur « traduction » en clair pour l'utilisateur.

Décembre 2001

5. XML : Extensible Markup Language ou Langage Extensible de Balisage.