# QUELQUES ÉLÉMENTS DE LINGUISTIQUE ET DE TRAITEMENT AUTOMATIQUE DES LANGUES POUR UNE MODÉLISATION DU DISCOURS MUSICAL

*Patrick Saint-Dizier*
IRIT-CNRS Toulouse,
stdizier@irit.fr

**RÉSUMÉ**

Dans ce bref article prospectif, nous proposons quelques éléments propres à la linguistique formelle et au traitement automatique des langues qui permettent de modéliser finement plusieurs éléments du discours musical. Nous introduisons les structures de traits typés pour représenter des thèmes. Nous introduisons alors quelques opérations majeures que l'on peut effectuer sur celles-ci dans le cadre de la musique. Nous présentons et adaptons ensuite le cadre de la pragma-dialectique afin de représenter les articulations discursives d'une oeuvre à plusieurs niveaux de détail. Nous concluons sur l'emploi de la plateforme <TextCoop> que nous avons développée pour l'analyse de ces structures musicales.

## 1. INTRODUCTION

Les relations entre les formalismes et les outils d'analyse de la langue naturelle et de la musique ont fait l'objet de plusieurs investigations dans le passé selon le point de vue de la perception (Katz et al 2009) (Jackendoff 2009), ou selon un point de vue syntaxique, en particulier génératif (Jackendoff et al 2005), (Lerdahl et al 1996)). Les différentes analyses montrent, à juste titre, que les formalismes et outils pertinents de part et d'autre sont relativement éloignés et que les analogies (il s'agit bien de deux langages) sont relativement éloignées ou abstraites. C'est particulièrement le cas, par exemple, de la syntaxe et de la morphologie en langue qui trouvent peu de contreparties directes en musique. C'est moins vrai en ce qui concerne l'analyse du discours : l'analyse du discours en langue et celui pratiqué par les spécialistes de l'analyse musicale ont en effet des similarités marquées.

Dans cet article, nous proposons d'abord une modélisation de thèmes (mélodies) par le biais de structures de traits typés largement utilisées dans les formalismes du traitement des langues. Nous montrons que des opérations assez simples permettent de rendre compte des variations apportées à un thème ainsi que des opérations classiques du contrepoint (formes miroir, rétrogrades, par augmentation, etc.). Les exemples étudiés dans cet article sont élémentaires, cependant ce formalisme permet aussi de représenter des structures et des processus plus complexes de la musique contemporaine. En utilisant ces structures typées, nous montrons à partir de deux oeuvres comment on peut développer une analyse du discours musical sur des niveaux de granularité variable (Peeters et al 2009). Le discours reflète la structure globale d'une oeuvre à travers les relations entre les thèmes et leurs développements. Pour ce faire, nous nous appuyons sur les développements récents de la théorie pragma-dialectique, une théorie linguistique du discours basée sur les interactions entre entités (locuteurs ou bien thèmes ou voix dans le discours musical). Enfin, nous montrons brièvement quelques éléments de modélisation et d'implémentation dans le cadre de notre plateforme d'analyse du discours, <TextCoop>, basée sur la logique et la notion de programme logique.

## 2. LES STRUCTURES DE TRAITS TYPÉS POUR REPRÉSENTER DES THÈMES

Les structures de traits typés sont largement répandues dans les formalismes de description de la structure grammaticale d'une langue. Elles sont à l'origine de plusieurs formalismes basés sur l'unification. Nous montrons ici comment un thème (pour simplifier ici une mélodie, mais toute structure peut être traitée de la même manière) se représente par une séquence de structures typées. Nous montrons ensuite comment des opérateurs se définissent sur une telle structure pour modéliser ou réaliser des opérations classiques : variation, passage au mode mineur, ou bien des opérations du contrepoint (formes miroir, etc.). La définition d'une structure correspondant à une mélodie élémentaire ou à un énoncé élémentaire a fait l'objet de plusieurs études, par exemple (Bimbot et al 2012), (Schauer 2006).

### 2.1. Quelques définitions

Une structure de traits typés est une structure récursive de la forme attribut-valeur où la valeur est un type, qui peut lui-même être une suite de paires attribut-valeur. Dans la langue naturelle, des attributs sont, par exemple, la catégorie syntaxique, la morphologie, qui se décompose en sous-attributs tels que genre, nombre, et voix, temps et mode pour les verbes. La catégorie syntaxique est associée à un ensemble de valeurs qui définissent son type : {nom, verbe, adjectif, préposition, ... }. Une telle structure se représente en général de la façon suivante pour un nom tel que *chaise* :

```
[cat= n,
 morphologie = [genre = féminin,
                nombre = singulier],
 type-sémantique = objet-physique ]
```

Les attributs peuvent être définis dans un ordre quelconque en respectant les hiérarchies. L'absence d'un attribut signifie dans notre cas qu'il n'est pas pertinent pour ce que l'on représente. Lorsque plusieurs objets sont décrits par une telle structure (par exemple les mots d'une phrase), il est possible d'exprimer des contraintes ou des relations entre ces objets ou bien de gérer ces structures par le biais d'équations, par exemple dans des règles de réécriture. On peut ainsi dire que le genre et le nombre d'un déterminant (ou article) sont identiques à celui du nom qu'il introduit. La même contrainte s'exerce entre le sujet et le verbe d'une proposition.

Les systèmes à base de traits ont fait l'objet d'une théorisation poussée dans un cadre logique (Carpenter 1992), ils sont aussi implémentés dans des langages logiques permettant la description déclarative et à un haut niveau d'abstraction de structures linguistiques. Ces structures peuvent alors être traitées par une grande variétés de stratégies d'analyse et d'algorithmes.

### 2.2. Représenter une mélodie par le biais d'une structure de traits

Nous restons ici dans un cadre volontairement très simple au niveau des exemples pour mieux introduire les concepts. Ce cadre peut être étendu à de nombreuses autres caractéristiques (par exemple le timbre ou l'accentuation) selon le même principe. Une mélodie, qui forme un thème, est une séquence de notes ordonnée temporellement, dans le cadre d'un mètre et d'une tonalité donnés.

Un certain nombre de caractéristiques d'une note peuvent être représentées indépendamment du reste du thème. Parmi ces caractéristiques, une structure d'attributs du type suivant (sans vouloir être exhaustif) peut être introduite :

```
[nom , hauteur, durée, ajout-durée,
   altération]
```

Le nom est simplement, dans notre exemple, à prendre dans la gamme de do à si, la hauteur est un nombre de 1 à 8 en général qui situe l'octave, la durée est plus complexe : au niveau du type, nous avons une structure ordonnée en progression arithmétique qui va de la ronde (ou la carrée) à la quadruple croche, les triolets ou autres regroupements sont indiqués de la même manière. L'attribut ajout-durée indique les extensions : point, double-point par exemple. Altération comprend le dièse, le bémol, double dièse et bémol et pas d'altération. La première note de la mélodie au soprano donnée en annexe 3 (op. 9 de Brahms) est le fa♯, il se représente comme suit, dans un format simple et lisible :

```
note: [nom = fa, hauteur = 5, durée = noire,
 sous-durée = aucune, altération = # ]
```

Cette notation peut paraitre simpliste par rapport à certains formats proposés par ailleurs et intégrés dans des notations de type XML, par exemple, (Tunrball et al 2008) ou (Peters et al 2009), ou des éditeurs tels que Lilypond. Toutefois, nous verrons que ce type de structure peut faire l'objet d'abstractions et d'opérations assez puissantes qui permettent, en analyse, de rendre compte de la structure musicale et du traitement des thèmes dans une oeuvre. Employée en génération, cette structure permet aussi de produire une grande diversité de possibilités.

Une mélodie intègre une suite de notes, chaque note étant représentée par le constructeur de type `note` comme illustré ci-dessus. Cette mélodie a elle aussi une structure de traits en propre qui se décompose en deux parties : (1) les caractéristiques globales à cette mélodie, par exemple le mètre, l'armature, le tempo ainsi que le numéro de mesure où elle commence dans l'oeuvre et (2) la séquence de notes. A partir de ces caractéristiques globales, en considérant la séquence de notes, on peut identifier par exemple les temps forts de la mélodie (sur la base du mètre et éventuellement du tempo : un 3/8 lent n'aura pas la même distribution qu'un 3/8 rapide), l'ambitus de la mélodie ou bien encore les modulations à partir de l'armature.

Le constructeur de type est `melodie`. Une structure d'attributs pour une mélodie est, par exemple :

```
melodie: [nom = (chaine de caractères libre)
 mètre , armature, tempo, mesure-début,
 liaison, nuance,
 seq-notes = [ note: .... ]]
```

La mélodie donnée en annexe 3 (variation 10 de l'op 9 de Brahms), a la structure de traits typée suivante pour ses deux premières mesures :

```
melodie: [nom = 'var10',
 mètre = 2/4 , nuance = piano,
 armature = 2#, tempo = 'poco adagio',
 mesure-début = 1,
 seq-notes = [
  note: [nom = fa,
  rang = 1, % position dans la mélodie
  hauteur = 5, durée = noire,
  sous-durée = aucune, altération = #],
  note: [nom = do, rang = 2,
  hauteur = 5, durée = noire,
  sous-durée = aucune, altération = #],
  note: [nom = ré, rang = 3,
  hauteur = 5, durée = noire,
  sous-durée = pointée,
  altération = aucune],
  note: [nom = si, rang = 4,
  hauteur = 5, durée = croche,
  sous-durée = aucune,
  altération = aucune ], etc.    ]]
```

On notera que lorsque l'armature est donnée, il est redondant d'indiquer les altérations pour chaque note, toutefois, en cas de modulation ou d'altération accidentelle, il est préférable de garder cette redondance. La séquence de notes est elle aussi réalisée indépendamment du mètre et du point de début de la mélodie. En cas de polyphonie, on ajoute au niveau de la mélodie la voix concernée. Les silences sont indiqués par la valeur 'silence' au lieu du nom de la note.

Enfin, le constructeur de type `melodie` permet de coder des informations qui portent sur des groupes de notes : liaisons de notes ou nuances par exemple. Si l'on veut décrire un crescendo de la note du rang 1 au rang 4 allant de la nuance *piano* à *mf*, l'attribut nuance reçoit alors un type plus complexe au lieu d'une valeur atomique :

```
melodie : [...
nuance = [ note:[rang =1, nuance = p]
           note:[rang =4, nuance = mf]
     intensite = crescendo]   ... ]
```

### 2.3. De la mélodie à la polyphonie

Dans le cadre d'une polyphonie (les aspects harmoniques ne sont pas traités ici par manque de place), les constructions présentées ci-dessus s'étendent à une oeuvre entière, horizontalement en traitant la ligne mélodique et verticalement en autorisant l'intégration de plusieurs voix.

Tout d'abord, une voix dans une oeuvre peut être considérée comme une suite de mélodies, avec des numéros de mesure de début qui suivent le déroulement dans le temps. Nous définissions alors le constructeur de type `voix` comme une suite de mélodies. Des relations de discours relient ces mélodies, elles sont présentées ci-après. Si nous reprenons l'exemple en annexe, nous avons :

```
voix :[ ident = soprano,
        mètre = 2/4 , nuance = piano,
        armature = 2#,
        tempo = 'poco adagio',
  seq-melodies:
    [melodie: [mesure-début = 1, etc.],
     melodie: [mesure-début = 4, etc.],
     ... ] ]
```

Les identificateurs tels que soprano peuvent être adaptés au type d'oeuvre. Le type `voix` hérite a priori des propriétés des types `mélodie` au niveau du mètre, de l'armature, de mesure-debut, etc. Toutefois, si une mélodie est dans un autre mètre ou ton, alors ceux-ci sont précisés à son niveau et l'héritage du niveau supérieur est bloqué. Ceci est une opération classique des structures de traits typés. Des contraintes d'intégrité ou de bonne formation portent sur la séquence de mélodies, par exemple, elles doivent s'enchainer et non se superposer, ceci est contrôlé sur l'attribut `mesure-début`.

Une polyphonie se définit par un ensemble d'au moins deux voix superposées. Le type `polyphonie` hérite d'attributs globaux : le mètre, l'armature et le tempo (a priori), mais pas la nuance qui peut varier selon les voix :

```
polyphonie :[ ident = 'nom oeuvre',
        mètre = 2/4 , armature = 2#,
        tempo = 'poco adagio',
seq-voix: [voix: [ident= soprano, etc. ],
           voix: [ident= alto, etc. ],
    ...  ] ]
```

Très globalement, et à titre d'exemple, on peut définir un canon à 2 voix à l'octave, avec un décalage de une mesure entre les voix en utilisant les attributs `mesure-debut` avec les valeurs 1 et 2 pour les deux voix si la voix supérieure commence la première et `hauteur` qui aura les valeurs 5 et 4 respectivement :

```
polyphonie :[ ident = canon2, mètre = 2/4 ,
     armature = 2#,tempo = 'poco adagio',
     seq-voix:[voix: [ident= 1,
     mesure-debut = 1, hauteur = 5, ...],
       voix: [ident= 1,mesure-debut = 2,
                hauteur = 4,...], ...]]
```

On note que les deux voix ont le même identificateur, ce qui permet de décrire un canon à un haut niveau d'abstraction, adéquat sur le plan de l'analyse musicale. Le formalisme de la réentrance peut être utilisé à ce niveau.

La représentation d'une mélodie ou d'une oeuvre, tout en demeurant assez lourde, offre un niveau d'abstraction intéressant avec les structures de traits. Cette approche permet par exemple de bien mettre en lumière les niveaux de structuration et les éléments partagés. Ces structures peuvent être représentées en XML. Le lien avec des représentations de plus 'bas niveau' comme MIDI, ou plus abstraites (celle de Lilypond par exemple) peut être défini mais reste à élaborer.

### 2.4. Les opérations génériques sur les structures de traits

Plusieurs opérations peuvent être définies de façon abstraite sur une structure de traits, par exemple pour développer des variantes à un thème. Nous en présentons quelques une ici. Ces opérations sont essentiellement des transformations apportées sur la valeur de certains attributs. Pour ce faire, il est nécessaire que ces valeurs soient organisées selon un principe mathématique. Par ailleurs, il est possible de définir des opérations sur ces structures typées à des fins d'analyse (par exemple repérer des modulations) et non de génération.

Les opérations présentées ci-dessous peuvent aussi bien être utilisées en production (produire de nouvelles formes) qu'en reconnaissance (donc analyser des formes existantes et reconnaitre les liens entre mélodies ou fragments de mélodies). Les règles décrites sont inversables. L'ensemble est codé en Dislog, un environnement de programmation en logique, développé sur notre plateforme <TextCoop> décrite très brièvement ci- dessous.

#### *2.4.1. La transformation par augmentation ou par diminution*

Cette opération, courante en musique baroque, consiste à augmenter ou à diminuer de façon proportionnelle la durée de toutes les notes d'une mélodie. La plupart du temps, on trouve une augmentation d'un facteur 2 (par exemple le premier thème de la fugue en ut♯ mineur du premier livre du clavecin bien tempéré). L'idée est de donner de l'emphase au thème. Dans les techniques de variation, on trouve aussi des diminutions, souvent de 50%. Dans les variations op 9 de Brahms étudiées ici (annexe 3), la variation 2 prend la basse du thème original comme thème

(c'est aussi le cas de la variation 10 évoquée ci-dessus, ce qui est très rare), chaque mesure du thème initial (en 2/4) est réduite à un temps (dans une mesure en 9/8) ce qui fait une diminution 25%. Enfin, dans la variation 10, les voix intermédiaires reprennent le thème en diminution (25% aussi) avec des notes intermédiaires, dans le style d'un contre-sujet d'un thème fugué).

Si l'on veut doubler les durées, il suffit de passer à la durée immédiatement supérieure (une noire devient une blanche, une note pointée devient doublement pointée, etc.). Le formalisme que nous utilisons (Dislog) offre un langage de spécification de ces transformations qui est déclaratif et permet de se concentrer sur le phénomène à traiter et non sur le calcul. Une règle pour l'augmentation par 2 s'écrit comme suit :

```
note: [durée = X, sous-durée = aucune],
[precede(X1,X)]  -->
note: [durée = X1, sous-durée = aucune].
```

Le prédicat `precede` définit X1, la durée immédiatement supérieure à X dans la suite ordonnée des durées. S'il y a une sous-durée, celle-ci est augmentée de la même façon.

### *2.4.2. Les formes miroir*

Les formes miroir sont fréquentes en musique baroque. Etant donnée une mélodie, la transformation consiste à inverser chaque intervalle : une tierce majeure ascendante devient une tierce majeure descendante. Les points de départ des deux formes (directe et miroir) peuvent être distincts. Des corrections peuvent être apportées par mutation si l'on veut préserver la tonalité. Le reste de la structure est préservé (les durées en particulier). Dans les polyphonies baroques, une forme miroir est superposée à sa forme directe (par exemple dans l'oeuvre d'orgue de Buxtehude et de Bach (ex. BWV 668, un choral figuré)). La variation 10 de l'op 9 de Brahms donnée en annexe commence par une mélodie au soprano et son miroir exact à la basse, le soprano commence par un fa♯ et la basse par un ré. Ce décalage permet de préserver la tonalité de ré majeur et d'assurer une harmonie stable.

La transformation se fait en plusieurs itérations en procédant par paire de notes. Un prédicat `miroir` calcule la note miroir. La transformation se code en Dislog comme suit (les noms de variables sont en majuscules) :

```
melodie: [nom = 'direct', ....,
 seq-notes = [ ...
  note : [rang = I, nom = N1,
          hauteur = H1, alteration = A1],
  note : [rang = J, nom = N2,
          hauteur = H2, alteration = A2],
... ]],
[J = I+1,
 miroir([N1,H1,A1,N2,H2,A2],
        [MN1,MH1,MA1,MN2,MH2,MA2])]
   -->
melodie: [nom = 'miroir', ....,
 seq-notes = [ ...
  note :[rang = I, nom = MN1,
         hauteur = MH1, alteration = MA1],
  note :[rang = J, nom = MN2,
         hauteur = MH2, alteration = MA2],
... ]].
```

Le prédicat miroir produit le miroir d'un intervalle, à partir du premier vecteur donné en paramètre.

### *2.4.3. Les formes inversées*

Les formes inversées consistent à jouer une mélodie dans l'ordre inverse des notes. En général, les durées sont assez similaires (par exemple, des croches comme dans l'Offrande Musicale) pour éviter des problèmes de métrique. Une forme directe peut être superposée à sa forme inversée. Sur la base des structures de traits présentées ici, il suffit d'inverser les rangs, le rang 1 devenant le rang N s'il y a N notes dans la mélodie.

### *2.4.4. Quelques autres transformations*

A partir d'une structure de traits, il est possible d'opérer un très grand nombre de transformations plus élaborées que celles présentées ci-dessus. Nous en évoquons quelques exemples ci- dessous.

Etant donnée une mélodie représentée par une structure de traits, il est possible d'y insérer ou d'y ôter des notes supplémentaires, en particulier des notes de passage ou des broderies variées. L'opération se fait sur la base d'une grammaire locale qui défini (dans Dislog) la structure des broderies et autres ornements. Sur les notes en existantes, cette grammaire va ajouter des notes et provoquer un décalage de rang et une réduction éventuelle des notes existantes en fonction de la longueur de la séquence que l'on veut insérer.

Il est aussi possible d'extraire un fragment d'une mélodie, entre les rangs I et J, pour en faire un élément autonome qui sera utilisé ailleurs. C'est par exemple le cas de fragments de thèmes utilisés dans des séquences de développement libre entre deux expositions de thèmes dans la forme sonate ou bien comme élément de divertissement entre deux expositions dans une fugue. Des variations de rythmes sont aussi simples à modéliser en modifiant la valeur de l'attribut durée.

L'alternance majeur-mineur peut aussi être traitée en (1) modifiant l'armature de la mélodie de façon adéquate et (2) en contrôlant les notes altérées dans la mélodie. De façon assez similaire, il est aussi possible de transposer une mélodie dans un autre ton.

Enfin, une analyse pas à pas de deux structures de traits correspondant à deux mélodies différentes, via les transformations évoquées ci-dessus, permet d'analyser les ressemblances et différences entre ces mélodies. Une analyse sémantique peut alors être opérée, par exemple (Barrington et al 2008).

## 3. POUVOIR D'EXPRESSION DE CE FORMALISME

Le système formel présenté ici, utilisé largement en syntaxe et sémantique formelle des langues, offre une grande puissance expressive. Une approche logique comme celle proposée dans le langage Dislog permet d'introduire de surcroit du raisonnement.

Nous n'avons pas à ce stade exploré toutes les possibilités expressives de ce système pour la musique, les exemples présentés ici sont demeurés très simples pour leur lisibilité. Il nous parait acquis que ces structures permettent de représenter des constructions musicales bien plus complexes et diversifiées que celles présentées. De ce point de vue, il nous parait possible de représenter les structures de la musique classique, mais aussi un certain nombre de structures 'discrètes' de la musique contemporaine. Par exemple, ce formalisme permet de décrire dans un type des **accords, des éléments d'une série, des clusters de notes, des réservoirs de sons** ou tout autre éléments global, ainsi que des formes de glissandi par exemple. La gestion temporelle des unités musicales peut être largement étendue, sans réference à un mètre a priori. Le modèle abstrait des structures de traits peut s'étendre à des types polymorphes ou des types sous-spécifiés, bien adaptés à certaines musiques comprenant une **composante aléatoire**.

Enfin, outre le caractère déclaratif des descriptions, la possibilité de décrire des transformations de façon déclarative et le recours à des contraintes de bonne formation ayant un caractère global en renforce la puissance aussi bien en analyse qu'en production.

Notre travail demeure essentiellement théorique. Il manque à celui-ci une interface élaborée comme celles développées dans PWGL et Open Music qui en donnent toute la facilité d'utilisation. Lilypond nous semble offrir essentiellement un langage d'édition. Il nous est un peu difficile d'opérer une comparaison sur la puissance ou la facilité d'utilisation des formalismes proposés dans chaque système du fait de leur différences en matière de cadre théorique. Le système de représentation que nous développons, basé sur des types, ainsi que les opérations 'logiques' et les transformations proposées nous semblent couvrir une large palette de phénomènes en écriture musicale. L'ensemble du système a été conçu pour permettre une analyse efficace du discours à un niveau d'abstraction adéquat. Sa proximité avec la langue nous semble être aussi un facteur intéressant, et pas seulement pour les musiques vocales, mais par exemple pour l'analyse du discours, des métaphores musicales ou bien des formes de la persuasion par la musique.

## 4. LE NIVEAU DISCURSIF

### 4.1. Les théories du discours en linguistique

Le niveau discursif est le moins exploré en traitement automatique des langues et en linguistique. Il semble aussi peu exploré en informatique musicale, alors qu'il est largement développé, mais de façon moins formalisée, en analyse musicale. Dans le domaine de la linguistique, la RST (Rhetorical Structure Theory (Mann et al 1988)) a été la théorie dominante pendant deux décennies. Elle a pour but de représenter les articulations des segments élémentaires (Elementary Discourse Units, EDUs) (Schauer 2006) d'un texte par le biais de relations de discours telles que : illustration, élaboration, justification, cause, conséquence, reformulation, etc. Si, à l'origine, seulement 22 relations ont été postulées, la littérature dénombre à présent environ 200 relations dont les contours sont parfois flous. La RST postule qu'une relation s'établit entre un noyau et un satellite (ou entre deux noyaux dans quelques cas). Par exemple, l'élément illustré (le noyau) est associé à son illustration (le satellite). Un texte est représenté par un graphe de telles relations. Une contrainte de bonne formation stipule que tous les segments de texte doivent être reliés les uns aux autres (Marcu 1997).

Plus récemment, la théorie pragma-dialectique (Van Eemeren et al 1984, 1992), associée à la théorie des actes de discours et à l'argumentation (Perelman 2012) (Moeschler 2002) a proposé un schéma d'analyse du discours basé sur l'interaction et l'argumentation. Cette théorie étudie les échanges selon une perspective descriptive (description des processus linguistiques) et normative (règles évaluant la validité ou la pertinence des énoncés). Cette théorie, en association avec les maximes de coopération de Grice, introduit 5 types d'actes illocutoires gérés par 10 règles du comportement dialogique et argumentatif dont la violation constitue un paralogisme. La théorie des actes de discours implicites (Walton et al 2008), (Reed et al 2007) développe ce point de vue dans des situations de dialogue naturel. Plusieurs relations sont introduites pour caractériser les relations entre locuteurs (apporter des preuves à un énoncé, contredire, etc.), ainsi que les relations entre les structures formelles des énoncés (inférence, contradiction, etc.).

Ces considérations, très brièvement exposées ici, peuvent paraître éloignées du discours musical. Ces théories se prêtent cependant très bien à la formalisation du discours musical via une transposition de leurs principes. De la RST, nous pouvons retenir le modèle noyau-satellite ainsi qu'un certain nombre de relations. Par exemple un thème (le noyau) peut faire l'objet de variations de différentes natures. Ces variations, considérées comme des satellites peuvent être caractérisées par des relations rhétoriques. De ce point de vue, la relation d'élaboration de la RST doit être considérée en musique comme une proto-relation qui doit être décomposée en plusieurs sous-relations. Les opérations d'augmentation ou de miroir peuvent aussi être considérées comme des relations rhétoriques.

Cependant, la RST modélise des relations qui demeurent statiques. Une oeuvre musicale ayant souvent une dimension argumentative ou rhétorique marquée (par exemple les sonates, les symphonies et les quatuors de Beethoven ou l'oeuvre d'orgue de Bach), il est important d'inscrire l'analyse du discours musical dans une perspective dyna-

mique de dialogue et d'argumentation.

Dans les pages qui suivent, et à travers des exemples, nous développons quelques éléments d'une approche d'analyse formelle du discours appliqué à la musique. Les segments du texte musical mis en relation que nous considérons ici correspondent au type `melodie` décrit dans la section précédente. Cette notion demeure vague quant à sa segmentation automatique dans une partition, c'est aussi le cas des EDUs en langue. Toutefois, l'analyse musicale sait parfaitement isoler des segments musicaux autonomes qui interagissent avec d'autres (par exemple avec des notations telles que A B A', A B A B A', etc. ou les structures de type da Capo, rondo, etc.). Ce sont ces segments qui seront mis en relation au niveau du discours musical. Pour des segments à forte articulation (par exemple avec des modulations importantes ou des fragments bien marqués), il est possible d'introduire la notion de CEDU (Complex EDU, segment élémentaire complexe) qui permet de caractériser plus finement des parties de ce segment qui seraient développées par la suite.

Il est aussi important de noter que notre analyse peut être horizontale (ou linéaire) ou verticale : si une mélodie est considérée globalement avec son harmonisation, alors les relations de discours seront établies entre segments qui se suivent, directement ou pas. Dans le cas d'une polyphonie, les relations peuvent être établies entre voix, si celles-ci ont une véritable autonomie, comme c'est le cas dans nombre de madrigaux ou, plus tardivement, dans les derniers quatuors à cordes de Beethoven (dont l'op. 135) ou de Bartok.

Les unités du discours musical, tout comme celles du discours en langue, s'identifient et se délimitent sur la base de marques. En langue, il s'agit essentiellement de connecteurs, de classes de verbes, de modaux, de la ponctuation et de la typographie. En musique, il s'agira par exemple de la structure des mélodies, de l'harmonie (cadences), des silences, du rythme ou du mètre. Sur cette base, et sans le développer dans ce bref article, il est possible de définir des règles de segmentation d'une oeuvre en unités élémentaires (voir aussi (Bimbot et al. 2010)). Le formalisme pour la segmentation est le même que celui présenté pour les transformations.

Nous développons à présent deux exemples qui illustrent notre propos. Il s'agit tout d'abord de l'exposition initiale des deux thèmes de la sonate pour piano 'Appassionata' op. 57 de Beethoven où nous avons un traitement linéaire du discours musical (Harding 2010). Nous présentons ensuite l'exemple donné en annexe 3, la variation 10 des variations op. 9 de Brahms sur un thème de Schumann, où est développée une analyse davantage verticale.

Nous donnons ci-dessous un bref commentaire écrit de chaque analyse. L'analyse formelle est donnée en annexe sous forme d'un schéma, comme il est de coutume en analyse des actes de discours en pragma-dialectique. Dans ces représentations, les blocs carrés à droite décrivent un segment musical d'une voix donnée sous sa forme externe, un numéro de voix, en cas de polyphonie, y est associée, les numéros des mesures sont indiqués entre parenthèses. Chaque segment est lié à sa représentation formelle (dans notre cas une structure de trait typée) via un arc qui décrit sa fonction illocutoire (par exemple : question, asserte, réponse, contextualise, soutien, harmonise). D'autres catégories sont en cours d'élaboration, en particulier celles liées à l'argumentation, où deux catégories émergent : support (par justification par exemple ou un autre opération rhétorique du même type) et attaque. Enfin, la force illocutoire n'est pas indiquée ici.

Les représentations formelles sont liées par des relations du type de celles présentées ci-dessus. Des relations peuvent co-apparaitre (par exemple tierce inférieure + miroir). Au niveau des segments musicaux originaux, une distinction est opérée entre éléments thématiques, notés A, B, A', B', etc. et les éléments non-thématiques, mais qui peuvent faire des emprunts aux thèmes, qui sont appelés transitions et notés Ti. Si une séquence est complexe, elle peut être décomposée, par exemple, dans le cas qui suit, A est décomposée en A1 et A2, car ces fragments sont appelés à évoluer indépendamment. Enfin, les transitions ont bien entendu une représentation formelle et souvent une fonction illocutoire.

### 4.2. L'exposition de la sonate op. 57 de Beethoven

Cette oeuvre très connue de Beethoven fut composée entre 1804 et 1806, deux ans avant la cinquième symphonie. Ces deux oeuvres ont des liens rhétoriques importants. C'est l'époque du testament de Heilingenstad, début d'un long débat de Beethoven sur le destin, doit-on l'accepter tel qu'il est ou se révolter. Dans plusieurs de ses oeuvres il en résulte un questionnement suivi de réponses, de batailles gagnées ou perdues. L'aboutissement, positif, se trouve dans ses deux dernières sonates, ses 6 derniers quatuors et dans le 9ème symphonie.

Le premier thème, en fa mineur, appelé A, débute à la mesure 1 (mesure complète hors anacrouse). Il dure 4 mesures, mais il est constitué de deux fragments de 2 mesures chacun A1 et A2. A1 est construit sur une cellule rythmique typique de 3 notes qui est répétée de nombreuses fois et qui sert aussi de motif dans les transitions. A est reproduit en A', en sol♭ majeur. A et A' ont la fonction illocutoire *question*, non seulement en raison du contexte de l'oeuvre, mais aussi et surtout en fonction de la forme du thème qui épouse l'intonation d'une expression interrogative en langue. Un tel type d'intonation est retrouvé à de multiples occasions dans l'oeuvre de Beethoven, par exemple, dans l'introduction lente de la sonate pathétique et de la 32ème sonate.

La transition T1 entre A et A' est 'vide' (silences). A2 est ensuite répété 2 fois. Suit une transition, T3, où apparait un fragment nouveau dans le registre grave du piano : le thème du destin, commun avec la 5ème symphonie. Cette transition va contextualiser la question posée par A et A' et nous ramener en fa mineur, ton initial. Suit A" (mesures 16 à 23), une amplification de A, avec une force illocutoire supérieure (batteries d'accords au lieu d'arpèges, ambitus plus large). Suit une transition, T5, puis le second thème B et sa variante B'. B est construit sur la même cel-

lule rythmique que A, mais avec une forme miroir, et un passage au relatif majeur, la♭ majeur. Nous avons donc ici un miroir et une alternance vers le majeur, renversement de la question en une réponse, ici très lumineuse, chantée presque comme un hymne. Cette analyse est présentée graphiquement en annexe 1 de ce document.

### 4.3. La variation 10 de l'op. 9 de Brahms

Les variations op. 9 sur un thème de Schumann sont dédiées à Clara Schumann, elle constitue en quelque sorte un témoignage de l'admiration et de l'affection du jeune Brahms pour le couple Schumann. Cette oeuvre date de 1854, elle coïncide avec l'internement de Schuman et reflète ces circonstances douloureuses. Cette série de variations, probablement l'une des plus belles oeuvres de Brahms, emprunte de nombreuses façons d'écrire à Schumann, et on y découvre les deux facettes de ce dernier, Eusébius et Florestan.

La variation 10 est la plus ample de ce cycle. Elle est la seule écrite en majeur (ré majeur), après 8 variations en fa♯ mineur et la neuvième en si mineur. Elle constitue une pause et un moment de détente dans cette oeuvre par ailleurs assez sombre. Dans cette variation, Brahms a voulu rendre hommage à Bach, celle-ci étant un vrai tour de force contrapuntique.

Cette variation a plusieurs éléments singuliers. Tout d' abord le thème n'est pas celui du thème initial, mais sa basse. Celle-ci est aussi utilisée en diminution comme thème de la variation 2. L'ambitus d'octave du thème traité dans cette variation a une valeur symbolique en musique baroque : l'octave représente allégoriquement le lien entre le ciel et la terre. Ici, avec les formes directe et miroir conjuguées, on a un double mouvement, du bas vers le haut (basse) et du haut vers le bas (soprano). La pièce est de plus peu modulante et donc très stable. Le thème (mesures 1 à 4) est d'abord exposé en même temps au soprano et à la basse en forme miroir tandis que les voix intermédiaires forment une sorte de réponse comme dans une fugue, réponse composée du thème, en diminution et à la tierce majeure supérieure. Les mesures 5 à 8 reproduisent à peu près le même schéma, en se terminant par une modulation en la majeur. De la mesure 9 à 12 (13), le thème et son miroir sont présentés au soprano et à l'alto, et en canon à une mesure, canon reproduit ensuite des mesures 13 à 16. A partir de la mesure 17, le canon se fait entre le soprano et la basse, mais de façon moins explicite pour cette dernière voix. L'oeuvre se termine par l'insertion du fameux thème de Clara, dans un passage plus modulant, où l'accompagnement en arpèges se fait à présent descendant dans le grave du piano.

Dans l'annexe 2 nous présentons des éléments de cette analyse dans le cadre de la pragma-dialectique.

## 5. BREF APERÇU DE <TEXTCOOP> ET DISLOG

Le formalisme des règles et ressources de <TextCoop>, DisLog (Saint-Dizier 2012), étend les possibilités expressives des approches à base d'expressions régulières et les adapte aux besoins de l'analyse de discours, en y intégrant des structures complexes (par exemple, les formes attribut-valeur) ainsi qu'un composant de raisonnement souvent utile dans l'analyse de telles structures. Notre approche s'appuie aussi sur les travaux, maintenant assez anciens, mais toujours actuels, des grammaires logiques (DCGs, XGs, MGs, etc.), qui s'appuient sur des modèles d'exécution inspirés des programmes logiques, dont Prolog. D'autres schémas de stratégie, par exemple à base de contraintes actives peuvent être associés. Cette approche nous parait intéressante pour l'analyse du discours, en langue ou en musique, en raison de son caractère déclaratif marqué, de son indépendance relative aux stratégies de traitement, et aussi de son aptitude à intégrer naturellement des modules de raisonnement et des structures de contraintes puissantes (par exemple des structures de traits typés, des contraintes d'arbres).

Le formalisme que nous proposons ici permet de coder des règles d'analyse de structures discursives conçues par des linguistes ou des musiciens et codées manuellement aussi bien que des règles issues de mécanismes d'apprentissage à partir de textes annotés, qui produisent en sortie des formes contraintes. Compte-tenu de la taille réduite de ce document, nous ne pouvons présenter ici que quelques éléments simples de <TextCoop>.

Le langage DisLog qui s'exécute sur <TextCoop> comprend les symboles suivants, ils suivent en général la syntaxe de Prolog et des DCGs (grammaires à clauses définies) :

- des symboles pré-terminaux et non terminaux. Les pré-terminaux se dérivent directement en des entrées lexicales ou des expressions (caractéristiques de structures rhétoriques ou de domaines, par exemple). Ces marques d'annotations peuvent faire référence à des structures déjà identifiées. Les symboles non terminaux font appel à des grammaires, essentiellement à caractère local. Les règles ne s'appellent pas entre-elles. Les liens entre structures sont réalisés par des opérations de liage sélectif. Les symboles non terminaux et préterminaux sont souvent associés à ou peuvent être directement des structures de traits attribut-valeur.
- des symboles terminaux indiqués entre crochets, cette possibilité est utile lorsqu'il y a peu de choix sur ces terminaux au sein d'une règle, dans le cas contraire, il est préférable de faire appel à un préterminal,
- des indications d'optionalité ou d'itérativité sur les symboles préterminaux et non terminaux,
- des symboles permettant d'exprimer la précédence linéaire,
- des 'gaps' qui représentent des séquences finies de termes qui ne présentent pas d'intérêt pour la règle

en cours de description. Les gaps peuvent être associés à des contraintes, en particulier des symboles terminaux ou non-terminaux qui ne doivent pas être ignorés.
- des appels à des prédicats qui introduisent des contraintes, des connaissances à intégrer ou des calculs divers.
- des fonctions d'assignation, explicites ou par défaut, d'étiquettes dédiées permettant d'étiqueter les structures reconnues avec d'éventuels attributs, calculés par les prédicats ci-dessus.

Les exemples donnés en section 2.4 montrent des emplois directs de Dislog sur des structures de traits typés. Une structure reconnue en partie gauche de règle est transformée, le résultat étant spécifié en partie droite de règle. Les appels à des prédicats Prolog sont intégrés juste avant le calcul du résultat. Pour l'heure, la plateforme TextCoop est dédiée au traitement de la langue, nous travaillons actuellement à une adaptation souple à l'analyse musicale.

## 6. PERSPECTIVES

Dans ce bref article, nous avons présenté plusieurs formalismes de la linguistique et du traitement du langage naturel qui peuvent être transposés assez directement à l'analyse musicale et à la formalisation du discours musical. Nous avons tout d'abord introduit les structures de traits typées, permettant de représenter des mélodies aussi bien que des oeuvres polyphoniques. Nous avons aussi développé quelques opérations qui s'appliquent habituellement sur des thèmes en montrant leur niveau d'abstraction et d'adéquation descriptive. Enfin, nous avons montré que le modèle des actes de discours et de la pragma-dialectique est adapté à la formalisation des articulations entre thèmes ou fragments thématiques dans le but de rendre compte de la structure et du fonctionnement rhétorique de l'ensemble d'une oeuvre.

Cet article conserve une dimension prospective. Il est clair que chaque dimension doit être approfondie, validée et testée. Les représentations que nous avons introduites et illustrées par des oeuvres du 19ème siècle sont utilisables et extensibles aux oeuvres antérieures (par exemple le baroque tardif où la symbolique des nombres est importante et doit apparaitre dans la représentation rhétorique) aussi bien qu'aux oeuvres plus récentes, où les notions des structures ont évolué à divers degrés (thème cyclique chez Franck ou Bartok, motifs en grappes chez Debussy, citations par exemple grégoriennes chez Tournemire, gestion des séries en musique sérielle, etc.). Enfin, comme rapidement évoqué, au niveau calcul, la plateforme <TextCoop> que nous avons développée pour l'analyse de la langue doit être adaptée, dès lors que les fonctionnalités utiles seront bien identifiées.

## 7. REFERENCES


[1] Barrington, L., Yazdani,M.,Turnbull, D.,Lanckriet, G., Combining Feature Kernels for Semantic Music Retrieval, ISMIR, pages 614-619, 2008.

[2] Bimbot, F.,Le Blouch, O.,Sargent, G.,Vincent, E. Decomposition into autonomous and comparable blocks : a structural decomposition of music pieces, Irisa research Report, 2010.

[3] Carpenter, R. The logic of typed feature structures, Cambridge university Press, 1992.

[4] van Eemeren, F.H., Grootendorst, R. *Speech acts in argumentative discussions : A theoretical model for the analysis of discussions directed towards solving conflicts of opinion*, Foris Publications, 1984.

[5] van Eemeren,F.H., Grootendorst, R. *Argumentation, communication, and fallacies : A pragma-dialectical perspective*, Lawrence Erlbaum Associates, 1992.

[6] Katz, J., Pesetzky, D. The Identity Thesis for Language and Music, research report, MIT, 2009.

[7] Harding, H. A. Analysis of Form in Beethoven's Sonatas, General Books, 2010.

[8] Jackendoff, R. Parallel and non parallel between language and music, *Music Perception*, vol. 26-3, 2009.

[9] Jackendoff, R., Lerdahl,F. The capacity for music : What is it, and what's special about it ?, *Cognition*, 2005.

[10] Lerdahl, F., Jackendoff, R. A Generative Theory of Tonal Music, MIT Press, 1996.

[11] Mann, W., Thompson, S. *Rhetorical Structure Theory : Towards a Functional Theory of Text Organisation*, TEXT 8 (3), pp. 243-281, 1988.

[12] Marcu, D. The Rhetorical Parsing of Natural Language Texts, ACL, 1997.

[13] Massin, J., Massin, B. *Beethoven*, Fayard, 1967.

[14] Moeschler, J. Speech act theory and the analysis of conversation, in Vandervecken D. et al. (eds), Essays in Speech Act Theory, Amsterdam, John Benjamins, 2002.

[15] Peeters, G., Deruty, E. Is Music Structure Annotation Multi-Dimensional ? A Proposal For Robust Local Music Annotation, LSAS, Graz (Austria), 2009.

[16] Saint-Dizier, P. Processing Natural Language Arguments with the <TextCoop> Platform, Journal of Argumentation and Computation, vol 3-1, 2012.

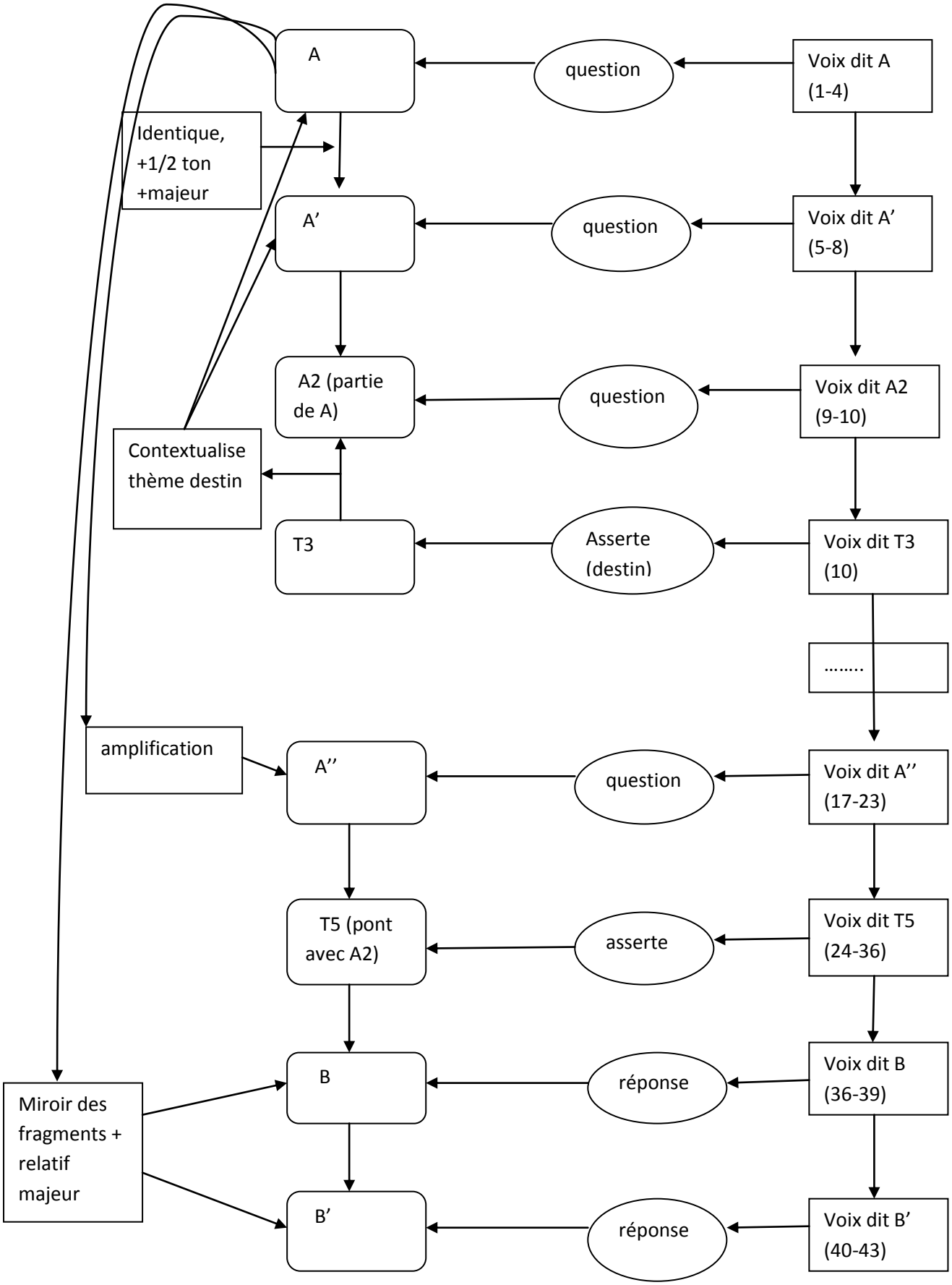
[17] Perelman, C. L'empire rhétorique : rhétorique et argumentation, Librairie philosophique Vrin, 2012.

[18] Reed, C., Grasso, F., *Recent Advances in Computational Models of Natural Argument* International Journal of Intelligent Systems,22(1), 2007.

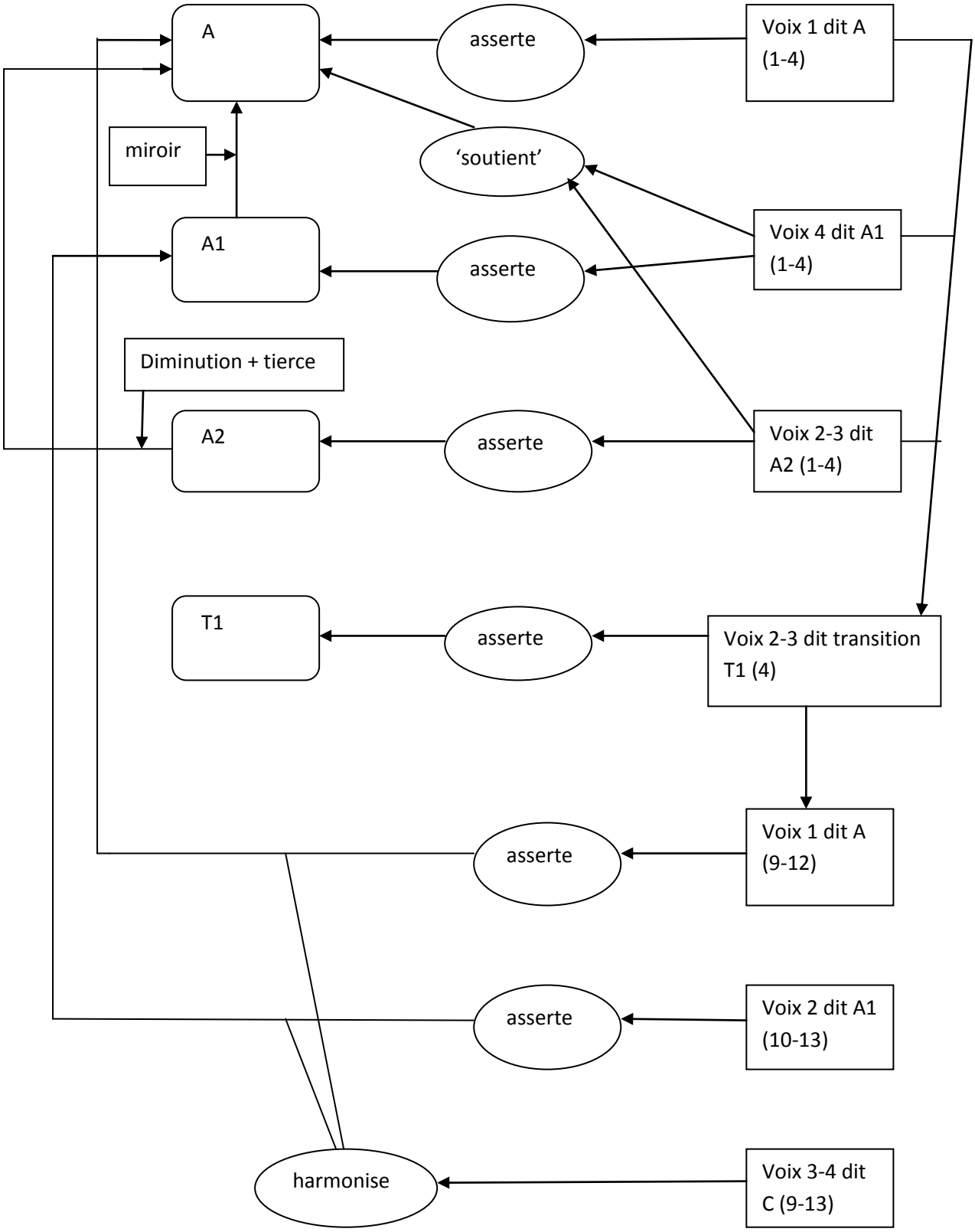
[19] Schauer, H. From elementary discourse units to complex ones, Sigdial workshop, ACL, 2006.

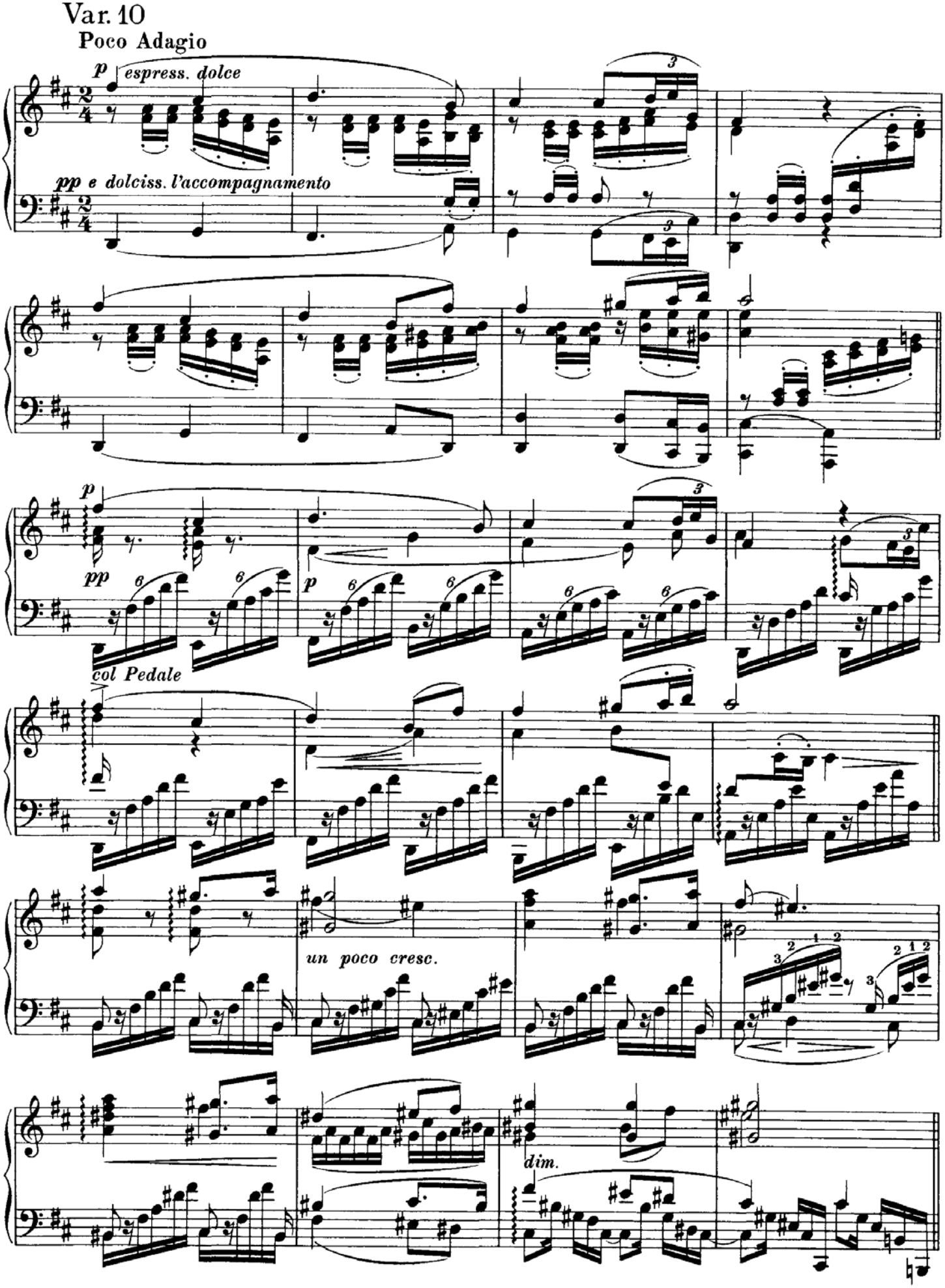
[20] Turnbull, D., Barrington, L., Lanckriet, G. Five Approaches to Collecting Tags for Music, ISMIR, pages 225-230, 2008.

[21] Walton, D., Reed, C., Macagno, F. *Argumentation Schemes*, Cambridge University Press, 2008.




**Annexe 1 : sonate op. 57 de Beethoven**




**Annexe 2 : var 10 de l'op. 9 de J. Brahms**

Var. 10
Poco Adagio
p espress. dolce
pp e dolciss. l'accompagnamento
col Pedale
un poco cresc.
dim.