

## A Taxonomy of Musical Genres

François Pachet  
Sony CSL-Paris,  
6, rue Amyot,  
75005 Paris, France  
E-mail: pachet@csl.sony.fr

Daniel Cazaly  
Sony Music France  
Paris, France

### Abstract

The recent progress of Electronic Music Distribution creates a natural pressure for fine-grained musical metadata. This metadata is needed to provide music distribution services which are able to cope with the mere size of music catalogues, and the desire of users to access music titles by similarity. In this context, we describe a project of a global music title metadatabase, and focus in the particular “genre” descriptor. We analyze existing taxonomies of musical genre as found in the music industry and on the Internet, and stress on their inconsistencies. We describe a novel music genre taxonomy based on a few guiding principles, and report on the process of building this taxonomy.

### 1. Introduction

Electronic Music Delivery (EMD) has recently benefited from technological progress in networking and signal processing. In particular, progress in networking transmission, compression of audio, and protection of digital data (Memon and Wong, 1998) allow now or in the near future to deliver quickly and safely music to users in a digital format through networks, either Internet, or digital audio broadcasting. Additionally, digitalization of data makes it possible today to transport information on content, and not only data itself, as exemplified by the Mpeg-7 standardization effort (Mpeg7, 1998). All these techniques give users, at home, access to huge catalogues of annotated music: a typical database of titles (e.g. Sony Music) contains about 500.000 titles (Amazon, Music Boulevard). A database containing all tonal music recordings would probably reach 4 millions titles. Adding ethnic music and non western types of music would probably double or triple this number. Every month, about 4000 CDs are created in Western countries.

The need for metadata describing the content of music catalogues has now become crucial, especially in the context of high level search services of electronic music distribution (Aigrain, 1999). However, good quality musical metadata for large catalogues is not yet available.

We have set up a project for designing and building such metadata for large catalogues of music titles. This metadata is intended to be used as an underlying knowledge layer for EMD services of music titles, such as music-on-demand systems, Digital Audio Broadcasting or Internet Radio. Each item of the database is described by a set of *descriptors* which take their value in a predefined taxonomy. The descriptors are of two sorts: technical and content descriptors. Technical attributes include the International Standard Recording Code (ISRC) when available, name of the title (e.g. “Learn to love you”), the name of the author (e.g. “Connick Harry Jr.”), the duration (e.g. “279 sec”), and the recording label (e.g. “Epic/Sony Music”). Content attributes describe musical properties of individual titles. The attributes include the following: *Genre* (e.g. “Jazz Crooner”) the subject-matter of this paper, but also *type of voice* (e.g. “muffled”), *music setup* (e.g. “instrumental”), *type of main instruments* (e.g. “brass”), *tempo* (e.g. “slow-fast”), and other optional attributes such as the type of melody (e.g. “consonant”), or the main theme of the lyrics (e.g. “love”). In the current state of our project, the database is created either by hand by experts (including the authors) for some descriptor or automatically for others (e.g. the tempo). This paper focuses on one particular descriptor of music titles: the musical genre.

## 2. Analysis of music genre classifications

The genre of music titles is probably the most important descriptor of all. It is widely used by both the music industry and the consumers. However, no sound classification of musical genre has yet been proposed. This section reviews the main genre classifications available.

### 2.1 Taxonomies of the Music Industry

The music industry has long created music taxonomies for its own needs. No effort - to our knowledge - has been done to unify these taxonomies, and indeed the task would be quite considerable, given the differences in design of these taxonomies.

The most important producers of music taxonomies are probably music retailers (e.g. Virgin Megastores, Fnac in France, etc.). Retailers produce taxonomies aimed at guiding consumers in shops, from the main entrance down to the record tracks. These taxonomies are usually made up of four levels: a main level with *global* musical categories (Classical, Jazz, Rock, etc.); a second level made of specific *sub categories* (e.g. "Concertos" within Classical, "Hard Rock" within "Rock"). The third level is usually an alphabetical ordering of *artists* (e.g. "Albinoni"). Within the artist rack lie the fourth level: artist *albums*. This kind of taxonomy has proved efficient for physical structuring of record shops. From the viewpoint of metadata, and hence automatic exploitation, however, they are largely inconsistent, since each level is designed to represent a different dimension: global music family for level 1, arbitrary sub categories for level 2, artists for level 3, and albums for level 4. Other music classifications can be found in large retailers, which are even cruder, e.g. by marketing categories (promotions/sales), or themes (e.g. sampler "rock collection", "best of love songs"), etc. In all cases, these classifications fulfill their role, which is to produce the shortest possible path for consumers to CDs, while keeping a reasonable CD rack size, and meaningful category headings.

Another inconsistency of these taxonomies is that artists produce music titles which may belong to slightly different musical categories within a given album. However, this is not relevant here: given the degree of generality of the subcategories, difference in genres in music titles usually remain consistent with large grained sub categories.

Music labels organize their catalogues in various ways, depending on the department using them. Although these classifications differ from the classifications of retailers, they keep the same overall artist and album based structure, usually inserting a "label" or "collection" layer.

Finally, copyright companies (e.g. ASCAP or SACEM) manage catalogues of music titles organized according to radio play list declaration or audience ranking, and are designed for copyright management. Therefore they include very crude taxonomies of genres (for instance soundtracks such as the Titanic are classified in Classical music category by copyright companies, because they are distributed by labels which usually produce Classical music (Sony Classical)).

### 2.2 Internet Taxonomies

Several taxonomies of musical genres have been designed by Internet music retailers and consequently made available to the public. These taxonomies are all aimed at helping users navigate in music catalogues using a top-down approach, like in record stores, but often with a much finer level of detail. Starting from main musical genres, such as "Classical", or "Rock", users progressively refine genres until they fall down on "leaf" nodes representing actual concrete genres, where they typically find lists of related albums or titles.

We have analyzed three of these classifications. The classification of Amazon, AllMusicGuide, and of the Mp3 Internet sites (a third important taxonomy, the one of CDDB<sup>2</sup> was not taken into account for time reasons. It includes 20 meta genres and 200 subgenres). From the structural viewpoint, each classification is organized as a hierarchy (i.e. no multiple inheritance links), and describes most of western music. Table 1 reports the number of genres for each taxonomy, and the number of so-called "meta genres", or top nodes in the hierarchy.

	Nb of genres	Nb of meta genres
AllMusicGuide	531	5
Amazon	719	18
MP3	430	16

**Table 1. Nb of nodes and root nodes for three largely used classifications of musical genre.**

An analysis of the relations between the taxonomies shows that: 70 words are common to the three taxonomies (e.g. “Texas Blues”, “Trance”, “Flamenco”); 198 are common to only two taxonomies, and 802 are specific to one taxonomy. This is due to the fact that the classifications clearly focus on specific genre areas against other. For instance the classification of AMG is very detailed in the Blues types, and contains 83 different Blues genres, some very detailed (e.g. “Modern Electric Chicago Blues”). Mp3 is on the opposite very detailed in the “Metal” area, but not so much in “Pop”, etc. The analysis shows clearly that there is not much consensus in these classifications, either from the lexical viewpoint (names used), and the structure (depth and structure of the hierarchies). Even largely words like “Rock” or “Pop” do not have common definitions ! Finally, it is not specified how new emerging genres (e.g. “Zouk-Love”) are taken into account.

The worst part is however the semantic viewpoint: it is not always clear what the hierarchical link between genres actually means. More precisely, we have identified the following possible meanings of the links in classifications (“A>B” denotes a hierarchical link from genre A to genre B):

- Most of the time, the hierarchical link is *genealogical*, i.e. denotes a musical evolution. For instance, “Pop” is a father of “Disco” (in Amazon). Of course, this genealogy is debatable, since musical evolution often involves several “father” genres. For instance, the emergence of “Disco” could also (arguably) be traced back to Soul and Pop.
- A *geographical inclusion*. For instance: “International”>“Africa”>“Algeria” in Amazon  
The presence of geographical information in genre taxonomies may in some cases make sense, but not always. It is quite debatable for instance that *Algerian* music is always stylistically different from, say, Tunisian Music.
- A mere *aggregation* (e.g. “R ’n B/Soul” >“Soul” and “R ’n B/Soul” >“R’nB” in AllMusicGuide),  
Aggregation is often used to put together genres that have nothing in common.
- A *repetition* (“Dance” > “Dance” in AllMusicGuide). This repetition often means that a given term (e.g. “Dance”) is polysemic, and denotes both a container (a generic genre) and a containee (an actual genre). Although this kind of ambiguity may be resolved in automatic system, the rule should however be made clear. In particular, it should be stated whether the generic genre can actually be used to describe music titles (in which case there would be an inconsistency in case of ambiguous terms) or if only leaf nodes can be used as descriptors.
- A *Historical period*. The “Classical” section of Amazon includes a node “Historical Periods”, itself containing “Baroque”, “Classical” (the musical period spanning from 1770 to 1830), “French Impressionist”, etc. Here again, chronology may be in some cases a good way to describe genre-subgenre relationships, but it then should probably be a different classification in itself, and not mixed with other dimensions, such as:
- *Specific dimensions of the subgenre*. For instance, the Classical section of Amazon, under “Classical Instrumental” > “Ballets and Dance” proposes a list of dance types: “Ballets”, “Ballettos”, “Basse Danse”, etc. Another node is “Instruments”, in which music is divided according to main instruments (“Brass”, “Keyboard”, etc.).

Of course, no consistency is maintained: most of the time, the classification oscillates between these different interpretations. In a context where the aim is to help users browse manually through a catalogue, this

inconsistency is not an issue, because users can interpret the links easily from a few examples. However, in a context where metadata is to be exploited by software to produce systematic search, some sort of consistency must be achieved.

### 3. Design of a new Genre Taxonomy

#### 3.1 Objectives

Based on the criticisms of existing taxonomies of musical genres, we have undertaken to design a new taxonomy of musical genre, spanning the whole domain of western music. By opposition with existing taxonomies, this taxonomy is designed to describe individual *music titles*, and not artists or albums. By convention, only terminal genres are used to describe music titles. More precisely, the objectives of this taxonomy are the following:

- The taxonomy is to be as *objective* as possible. Each taxon should have definitions in terms of objective criterion, and should not be strictly lexical, i.e. assume shared vocabulary conventions.
- The genre descriptor should be *independent* from other descriptors in the metadatabase. This independency of descriptors is a fundamental notion in the design of the metadatabase, and applies to all descriptors. However, it apply particularly well to the genre descriptor as we will illustrate.
- The descriptors should be designed to support search by *similarity*. Similarity is the main constraint to decide whether genre groups should be created or not.
- The taxonomy should be *consistent*. The semantics of the link should remain the same throughout the taxonomy, and support *evolutivity*, i.e. be able to cope with new emerging genres.

#### 3.2 Design Principles

To describe the design of our taxonomy we first have to recall that genre is only one of the many descriptors of the metadatabase. Current other descriptors include: tempo, rhythm type, voice type, main instruments, etc.

In this context, we addressed the four objectives listed above as follows:

- *Objectivity*. One way to achieve some sort of objectivity is to define the specifics of each genre using 1) a differentialist approach and 2) a reference defined in terms of orthogonal descriptors.
- *Independency*. Independence of descriptors led us to 1) reconsider “obvious” genre taxons as inappropriate, and 2) create new, orthogonal descriptors.
- *Similarity*. The need for genre taxons to support similarity links was the driving force to decide whether some taxons should be reified or not (examples ?). Two relations were therefore designed: a strictly hierarchical relation indicating stylistic inclusion, and an explicit similarity relation establishing links across stylistic regions.
- *Consistency/Evolutivity*. These principles were applied uniformly. This does not imply that our hierarchical link does not, in some cases, parallel other links such as geography (e.g. “Rock:California”, which has a specificity of its own in terms of instruments).

We will review now each of these principles and illustrate it with typical examples.

#### 3.3 Objectivity

Objectivity is a very difficult notion when talking about music genres. In our context, grounding of music genres (e.g. seeing genre as a feature to be extracted from the signal) is clearly impossible. The only thing we can target is to ensure that the structure of the taxonomy relies on objective arguments, rather than on intuitions derived from taxon names. To achieve this aim, we use two principles: differentialism and inter descriptor reference. Additionally, we document each terminal taxon with examples of music titles.

### 3.3.1 Differentialism

The notion of differentialism stems from the field of linguistics (Rastier et al., 1994), in which researchers have long battled to find reasonable definitions of semantics, and cope with the delicate problem of defining lexical terms through a reference to the “world”. Differentialism is an approach in which the meaning of terms is not given by an external referential (some objective element of the world), but simply by a description of how the taxon differs from other taxons of the base. Following this inspiration, Bachimont (2000) systematizes the approach and proposes guidelines for building hierarchical taxonomies (in the medical domain). These guidelines consists in explicitly stating, for each taxon, how it relates with its close neighbors: father and siblings. This relation is stated both by explicit similarities and explicit differences. In our case, we choose to state explicitly only differences with the father, and differences with sibling genres.

### 3.3.2 Inter Descriptor Reference

To state differences between genres (either genre to father, or genre to sibling), we need a language. To avoid introducing yet another degree of complexity and fuzziness in the database, we choose to use *other descriptors* of the database as a description language. This hypothesis is actually very strong, as it implies that the other descriptors cover the music domain with a sufficient degree of detail to allow the characterization of subgenre differences. However, our experiments show that it is a very good hypothesis, which indeed helps in structuring the genre hierarchy. We give here some examples of genre/subgenre using this approach.

Here are examples of stylistic differences related to differences in instruments:

Genre: **Blues**

Father: none (root genre)

Examples: Non terminal

Genre: **Rhythm and Blues**

Father: Blues

Difference with father:

Instrument Descriptor:

- massive presence of *Brass* instruments

Examples: Non terminal

Genre: **Funky Music**

Father: Rhythm and Blues

Difference with father:

Instrument Descriptor:

- massive presence of *funk guitar*, and *Bass*

Examples: Sex machine / Brown, James; I want you back / Jackson 5, The

Genre: **Soul**

Father: Rhythm and Blues

Difference with father:

Instrument Descriptor:

- massive presence of *String*

Voice type Descriptor:

- *soft* voices

Examples: Non terminal

Genre: **FM**

Father: Rock

Difference with father:

Instrument Descriptor:

- massive presence of *Synthesizers*

Examples: Hold the Line /Toto ; Eye of the Tiger / Survivor

Genre: **California**

Father: Rock

Difference with father:

Instrument Descriptor:

- massive presence of *12-String guitar*
- less presence of electric guitar

Examples: It's a heartache / Tyler, Bonnie; Maggie May / Stewart, Rod

Here are examples of stylistic differences related to a difference in rhythm:

Genre: **Mento**

Father: Rhythm and Blues

Difference with father:

Rhythm Descriptor:

- Calypso

Examples: Non Terminal

Genre: **Ska**

Father: Mento

Difference with father:

Tempo Descriptor:

- Faster

Instrument Descriptor:

- Presence of *Brass section*

Examples: Guns Of Navarone / Skatalites, The; Al Capone / Prince Buster

### 3.4 Independence

The independence criterion has led to some surprising decisions, regarding some well known term, which in fact were not appropriate. This rule has led us for instance to give up including "Rap" from the taxonomy of genres, considering that rap is more a diction type than an actual stylistic specificity. An advantage of having Rap in an independent criteria is that this avoids having to create numerous "Rap" extensions of existing genres (e.g. Rock>Rap, Reggae>Rap), etc.

One of the most problematic term was the so-called *Variété* and *Easy Listening* music genres. These categories concerns a vast area of popular music, which is very difficult to describe objectively. For instance, artists like Georges Brassens (well known French song maker) would not be considered as *Variété*, whereas Johnny Hallyday (well known French Rock singer) would be. Similarly to the "Rap" problem, *Variété* is something that can occur in different areas of the genre taxonomy (US *Variété*, also called "International" in record stores, but also Italian *Variété*, etc.). Instead, an analysis of the failure in providing distinctive musical features of *Variété* and *Easy Listening* led us to introduce a specific descriptor: the audience location. Audience

describes the typical place where the music is usually listened to. The main values for this descriptor are: "Popular" (i.e. played on generalist radios), "Specialized" (played on specialized radios), "Discotheque" (played mostly in discotheques), "Ballroom Dancing" (frequent location of Variété music), and "Easy Listening" (public places). This descriptor has the advantage of being objective (one can get easily radio play lists for instance), and also avoids having to introduce ranking in music styles (Variété or Easy Listening being often considered a "minor" style by some people). Following the same approach, we introduced a descriptor to describe the *danceability* of certain music titles. Danceability describes what dance type (if any), the music suggest, and can take various values such as "no" / "rock" / "salsa", etc. This descriptor allows to get rid of numerous extensions of genres ("Rock>Dance", "Pop>Dance"), while providing a way of differentiating between titles in a given category. For instance, *My Sharona / Knack, The* is a "Pop>Garage" title, which is not danceable. Conversely, *Wild Things / Troggs, The* is a danceable "Pop>Garage" title.

Similarly, we gave up on a certain number of terms which would include orthogonal descriptors such as "Symphonic Pop" (an existing term in existing Internet classifications), which indicates a difference in the orchestration (Symphonic band instead of standard Rock band).

### 3.5 Similarity

The similarity relation of our taxonomy is a symmetric binary relation, used to establish links across stylistic regions. Following the differentialist approach, we document each of these links by stating explicitly what are the differences and similarities of two linked genres in terms of other descriptors.

For instance "Rhythm and Blues>Funky Music" is related to "Funk>Electro" with the argument that they both have a massive presence of *Electric Bass*. "Rhythm and Blues>Honky Tonk" is related to "Rock>Hard Boogie" because they share the same rhythm type (*Boogie*); "Rhythm and Blues>Tamla Motown" and "Soul>Disco>Philadelphia" have the same orchestration (*Brass* and *Strings*).

These similarity relations are the most important information of the taxonomy since they will eventually be used to establish similarities global between titles.

### 3.6 Consistency/ Evolutivity

The hierarchical link may be seen as a representation of the genealogy of music genres, although it is not its primary aim. This genealogy is taken into account only superficially. In particular when an evolution is multiple (which is often the case), we choose a main father genre. In practice, there is always a consensus for such a main root genre. For instance, Disco is represented as a descendent of Soul in our taxonomy. This raises in turn a question regarding the evolution of the database: what happens when a new genre is emerging ? The solution we adopt is to create systematically when needed a subgenre called "Roots", as soon as a genre becomes non terminal. This subgenre (always terminal) is then used to describe titles whose genre belongs to the original genre "before it started to expand". For instance, some titles such as *Try a little tenderness / Otis Redding* are described as "Rhythm and Blues>Roots" because they belong to the genre "Rhythm and Blues" before it started to create subgenres (such as *Funky Music* or *Soul*). For the same reason, *I'm Gonna Love You Just A Little More, Baby / Barry White* are "*What's going on / Gaye, Marvin*" considered "Soul>Roots".

## 4. Conclusion

Our taxonomy currently contains 378 genres and about 800 similarity relations. This paper reported on the process of building the taxonomy and stressed on the consequences of applying a few guidelines as systematically as possible. Most notable are the disappearance of well known terms (such as *Rap* or *Variété*), and the need for introducing orthogonal descriptors such as Audience location and danceability.

Of course, the decision to reify a term and include it in the taxonomy can only be based on a consensus. The independence criteria and the objectivity criteria somehow contradict each other: when a new genre emerges, it is often, if not always, by the addition of some element to an existing genre (definition of objectivity). Whether or not this addition should be represented as a difference in other descriptors, or by adding an explicit new genre, is

based on assessing the relative importance of the genre within the taxonomy, and the interest of the new genre in terms of the similarity relations which can be introduced, which remains very much a subjective process.

However, the guidelines presented here were used to produce a taxonomy of genres, used to annotate 5000 titles, with a target of 10.000 in the near future (about 25 titles per genre). The taxonomy is by no means universal, but is to be seen as a starting point for developing high level content-based automatic music search systems. We have already used it in various EMD projects, including an automatic music program scheduler (Pachet et al., 1999). Current work concerns focus on the scale-up of the taxonomy and the metadatabase (using automated tools such as Protégé (Grosso et al., 1999) and its effective use for building EMD services.

## 5. References

Amazon web site: <http://www.amazon.com>

All Music Guide web site: <http://allMusic.com>

Mp3 web site: <http://www.mp3.com>

Aigrain, Philippe (1999) "The new applications of content processing of music", *Journal of New Music Research*, 28(4).

Bachimont, Bruno "Engagement sémantique et engagement ontologique: conception et réalisation d'ontologies en ingénierie des connaissances", *Ingénierie des connaissances*, Charlet, Jean Eds, INRIA, 2000.

Grosso, W. E. Eriksson, H. Ferguson, R. W. Gennari, J. H. Tu, S. W. & Musen, M. A. Knowledge Modeling at the Millennium (The Design and Evolution of Protege-2000). Stanford Medical Informatics report 1999-0801 1999.

Mpeg-7 (1998) Context and objectives, International Organization for Standardization, report ISO/IEC JTC1/SC29/WG11, Oct. 1998

Memon, N, Wong, P. W. (1998) Protecting Digital Media Content. *Communications of the ACM*, July 1998, pp. 34-43.

Pachet, F. Roy, P. Cazaly. D. (2000) "A Combinatorial approach to content-based music selection". Proceedings of the IEEE Multimedia Computing and Systems Int. Conf., Firenze, (Italy), pp. 457-462. To appear in IEEE Multimedia, March 2000.

Rastier, F. Cavazza, M. Abeillé, A. (1994) *Sémantique pour l'analyse*, Masson, Paris.