



Projecte del treball de final de grau

Les emocions en la música: estudi de
l'efecte de la veu, la instrumentació i la
producció en el reconeixement
d'emocions

Estudiant: Víctor Núñez Tarifa
Especialitat/Àmbit/Modalitat: Sonologia
Director: Perfecto José Herrera Boyer
Promoció: 2013-17

Índex

Agraïments

Resum

1.- Introducció	1
2.- Estat de la qüestió	3
2.1.- Què és l'emoció?	3
2.2.- Principals formes d'emoció	4
2.3.- Com podem mesurar les emocions?	9
2.4.- Quina relació tenen la música i les emocions?	11
2.5.- Com responen els oients a la música (percepció vs inducció)?	14
2.6.- El reconeixement d'emocions en la música	16
2.7.- Els descriptors d'àudio	18
2.8.- Les tècniques d'aprenentatge automàtic	21
2.9.- La producció musical com a creadora d'emocions	25
3.- Objectius	27
4.- Metodologia	28
4.1.- Creació de la col·lecció d'àudios	28
4.2.- Avaluació d'emocions i extracció d'annotadors a partir de la col·lecció	30
4.3.- Anàlisi estadístic descriptiu de les dades obtingudes en el qüestionari	37

4.4.- Extracció de descriptors per cadascun dels arxius musicals de la col.lecció	38
4.5.- Creació de models	41
4.6.- Reproductibilitat	42
5.- Resultats	44
5.1.- Anàlisi estadístic descriptiu de les anotacions	44
5.1.1.- Anàlisi dels subjectes	44
5.1.2.- Anàlisi de les emocions sense diferenciar condicions	48
5.1.3.- Anàlisi de les emocions diferenciant condicions	51
5.2.- Anàlisi dels resultats de la creació de models	58
5.2.1.- Anàlisi comparatiu del grau d'efectivitat dels models	58
5.2.2.- Anàlisi de les emocions sense diferenciar condicions	61
5.2.3.- Anàlisi de les emocions diferenciant condicions	62
5.2.4.- Anàlisi dels resultats de la combinació de condicions	65
6.- Discussió	68
7.- Conclusions	73
7.1. Contribucions	73
8.- Referències	75

Agraïments

Un especial agraïment al meu tutor del projecte Perfecto José Herrera Boyer, qui m'ha ajudat, prestat molta atenció i estimulat en al llarg del projecte.

També a la meva família, a Eliana Pedon i a en Jordi Bosch Molina per el suport incondicional que m'han mostrat en tot el procés.

Resum

En aquest treball abordem la caracterització i classificació d'emocions expressades en la música. Ho fem dins d'un context sonològic, ja que ens preguntem si aquesta comunicació d'emocions és diferent quan escoltem diferents materials dels que s'integren dins d'una cançó (veu sola, instrumentació, veu i instrumentació) a quan escoltem la peça completament produïda.

A partir de la revisió de la literatura sobre el tema, proporcionem una visió general de les característiques i problemàtiques més destacables en la classificació d'emocions expressades en la música. Per a tal fi es crea una col·lecció de fragments musicals d'estils i estímuls emocionals variats, dels quals s'extreuen quatre variants diferents: veu, instrumental, veu + instrumental i mescla; cadascun d'aquests fragments són avaluats per diferents subjectes en relació a les emocions percebudes en cada fragment. Les respostes són analitzades estadísticament i s'estableixen relacions entre els fragments i les emocions, tant a nivell de la predominança d'una emoció com de la relació entre elles en cadascun dels fragments.

En aquest sentit, s'ha pogut observar com, d'una banda, les emocions positives (alegria, animació) tendeixen a obtenir resultats de percepció més alts que les emocions negatives (irritació, despreci) i d'una altra banda, com els diferents elements que integren una cançó tenen un efecte rellevant en la percepció de les emocions.

A continuació es realitza un anàlisi acústic i s'extreuen els descriptors i característiques musicals més rellevants de cadascun d'ells per a ser utilitzats en un model que permetin predir quines seran les emocions percebudes pels oients en altres fragments i per a les diferents condicions musicals abans esmentades.

En aquest sentit, les emocions negatives han obtingut els millors valors de predicció, principalment en la condició mescla, la qual ha vist incrementada la seva efectivitat de predicció en la combinació amb els elements de les condicions veu i instrumental.

Resumen

En este trabajo abordamos la caracterización y clasificación de emociones expresadas en la música. Lo hacemos desde un contexto sonológico, ya que nos preguntamos si esta comunicación de emociones es diferente cuando escuchamos diferentes materiales de los que se integran en una canción (voz sola, instrumentación, voz e instrumentación) a cuando escuchamos la pieza completamente producida.

A partir de la revisión de la literatura sobre el tema, proporcionamos una visión general de las características y problemáticas más destacadas en la clasificación de emociones expresadas en la música. Para tal fin se crea una colección de fragmentos musicales de estilos y estímulos emocionales variados, de los cuáles se extraen cuatro variantes diferentes: voz, instrumental, voz + instrumental y mezcla; cada uno de estos fragmentos son evaluados por diferentes sujetos en relación a las emociones percibidas en cada fragmento. Las respuestas son analizadas estadísticamente y se establecen relaciones entre los fragmentos y las emociones, tanto a nivel de predominio de una emoción como de la relación entre ellas en cada uno de los fragmentos.

En este sentido, se ha podido observar como por un lado, las emociones positivas (alegría, animación) tienden a obtener resultados de percepción más altos que las emociones negativas (irritación, desprecio), y por otro lado, como los diferentes elementos que integran una canción tienen un efecto relevante en la percepción de las emociones.

A continuación se realiza un análisis acústico de cada uno de los fragmentos y se extraen los descriptores y características musicales más relevantes de cada uno de ellos para ser utilizados que permitan predecir qué emociones serán percibidas por los oyentes en otros fragmentos y para las diferentes condiciones musicales nombradas anteriormente.

En este sentido, las emociones negativas han obtenido los mejores valores de predicción, principalmente en la condición mezcla, la cuál ha visto incrementada su efectividad de predicción en la combinación con los elementos de las condiciones voz e instrumental.

Abstract

The project deals with the characterization and the classification of the emotions expressed by music. The analysis has been done from a sonologic perspective, considering that our aim is trying to discover if these emotions are differently perceived when we listen the different part of a song separately (voice, instrumentation, voice and instrumentation together) in respect of listening the whole song produced.

After reviewing the existent literature, we have focused our attention/have started the work introducing a global view on the most interesting features and issues of the categorization of the emotions expressed by music. With this aim, we have created a collection of soundtracks of different music styles able to transmit different emotional reactions/stimulus, by whom we have extracted 4 different versions: voice, instrumentation, voice + instrumentation and a mix of them; the emotions transmitted by any soundtracks have been analysed through a survey that we presented to a varied public.

We have statistically examined the answers and we have established connections between any soundtrack and the emotions perceived, considering the existence of a predominant emotion and the relationships between them in any soundtrack.

In this way, we have observed that positive emotions (happiness, animation) are more frequently perceived than negative (irritation, contempt); on the other hand, we have noticed that the different elements of a song have strongly impact on the emotions' perception.

Next, we have realised an acoustic analysis of any soundtrack and we have deduced the most relevant musical features of any one that allow to predict which emotions will be perceived by the listeners appreciating other soundtracks and under the different musical conditions mentioned before.

In this way, negative emotions have obtained best prediction values, mainly in mix condition; furthermore, emotions perception in mix condition has increased their values in combination with voice and instrumental features.

1.- Introducció

Les emocions són un element intrínsec en l'esser humà que defineix en molts casos les nostres pautes de comportament i es manifesta en múltiples ocasions en el nostre dia a dia, de formes molt diferents i per motius molt diversos;

En l'últim segle, la psicologia ha dedicat gran part dels seus esforços en donar resposta als processos i mecanismes que intervenen en l'estimulació d'emocions, així com a les principals formes en que aquestes es manifesten, i malgrat els avenços obtinguts, encara hi ha moltes preguntes a respondre.

Un dels àmbits on les emocions es manifesten de forma molt clara i diversa és en la música, la qual mitjançant les seves característiques sonores i acústiques és capaç d'evocar i induir emocions en els oients en situacions i contextos diferents (p. ex. l'escolta d'un fragment musical, la interpretació o composició d'una peça).

Des de l'àmbit de la psicologia musical s'ha posat molt d'interès envers la relació de les emocions amb la música, donant lloc a propostes de recerca que aporten coneixements nous en aquest camp, a la vegada que serveixen com a base per l'aplicació en altres àmbits, com per exemple el de la intel·ligència artificial (IA).

La IA ha estat usada en un nombre de camps molt ampli, incloent diagnòstic mèdic, comerç amb accions, control robòtic, lleis, percepció remota, descobriments científics, jocs i aplicacions musicals; d'aquest últim n'han sorgit propostes com l'emulació de l'estil compositiu de Johann Sebastian Bach (Cope, 2001), el disseny d'algorismes de recomanació de cançons utilitzats en programes de música en streaming (Celma, 2010) o la creació d'aplicacions capaces de detectar i modelar característiques pròpies de la música (Bogdanov et al., 2013).

Fruit del modelatge de les característiques pròpies de la música neix el present projecte, el qual pretèn estudiar l'efecte de les parts que componen una cançó (com poden ser la veu, la part instrumental o fins i tot la producció) en el reconeixement automàtic d'emocions, al mateix temps que recull informació envers quines emocions es perceben amb més claredat, quins components d'una cançó ajuden a aquesta percepció i en quin grau cadascun d'aquests components hi participen.

La idea d'aquest projecte sorgeix fruit de l'experiència personal com a sonòleg i guitarrista clàssic, en la qual se m'ha presentat de forma reiterada la inquietud de com poder fer arribar a l'oient unes emocions concretes i definides a partir de la

interpretació, la producció musical o la intel·ligència artificial.

Aquest estudi es realitzarà des d'una vessant musical i tecnològica, utilitzant tots els coneixements assolits en la formació sonològica que ajudin a analitzar les dades i a prendre les decisions pertinents.

Els resultats poden ser utilitzats en aplicacions futures, tant sigui per obtenir models de reconeixement d'emocions més eficients realitzant combinacions entre els components de les cançons que donin millors resultats, com per el desenvolupament d'eines de suport per a la presa de decisions en la gravació i producció de les cançons d'un disc.

2.- Estat de la qüestió

2.1.- Què és l'emoció?

Les emocions són un aspecte molt important en el comportament humà que es presenta en cadascun dels processos d'acció, percepció, memòria, aprenentatge i presa de decisions i es pot manifestar de formes molt diverses en el nostre dia a dia (p.ex. en una conversa amb un company de classe, en la rebuda d'una notícia agradable, veient una pel·lícula, un quadre, escoltant música o interpretant una peça en un concert). Però què es realment una emoció?

Aquesta pregunta, que per cert va ser el títol d'un dels treballs d'investigació més famosos del filòsof William James (1884), encara no ha rebut una resposta definitiva. Hi ha diverses raons que expliquen la situació en la que ens trobem. Una d'elles és que les emocions són difícils de definir i mesurar. Com van assenyalar Fehr i Russell (1984, p.464), tothom sap el que és una emoció, fins que es demana una definició per aquesta. Una altra raó pot ser el paper *perjudicial* atribuït a les emocions com a elements motivadors de la conducta humana. Tot i que el nostre punt de vista ha variat considerablement al llarg de la història, un dels punts de vista predominants al llarg d'aquesta ha estat el pensament impulsat per Descartes de l'emoció com a element *segrestador* de la raó (Cacioppo & Gardner 1999). Només recentment ha estat quan els investigadors han reconsiderat el paper exercit per les emocions en la racionalitat humana (Damasio 1994; De Souza 1987; Johnson-Laird & Oatley 1992; Simon 1967), el que ha portat al concepte *d'intel·ligència emocional* (Goleman 1995), tant de moda en l'actualitat i que havia estat inconcebible durant bona part del segle XX.

Durant els últims 100 anys, els teòrics han proposat una sèrie de definicions diferents de l'emoció. Kleinginna i Kleinginna (1981), en un intent de donar una visió més clara a la definició d'emoció, van trobar, a partir del recull de llibres de text, diccionaris i altres fonts, 92 acepcions diferents al mateix terme. Prenent com a referència totes aquestes acepcions, van proposar la següent definició consensuada (Kleinginna & Kleinginna 1981, p.355):

L'emoció resulta de la intervenció d'un conjunt complex d'interaccions entre factors subjectius i objectius a través dels sistemes neuronals/hormonals, que poden (a) donar lloc a experiències afectives com ara sensacions d'excitació, plaer/desgrat; (b) generar processos cognitius tals com efectes perceptius rellevants, avaluació d'events (appraisal), processos d'etiquetatge; (c) activar ajustos fisiològics generalitzats en reacció a certes condicions, i (d) portar a un comportament que sovint, però no sempre, és expressiu, amb un objectiu concret i adaptat.

A molts investigadors els agrada pensar en l'emoció com una seqüència d'esdeveniments: les emocions es *disparen* a partir de les avaluacions cognitives dels nostres esdeveniments més significatius; aquestes avaluacions evoquen fortes reaccions de la majoria dels nostres sistemes corporals; les reaccions corporals generen experiències subjectives de la sensació; i les experiències subjectives porten a l'acció i la conducta expressiva. No obstant, els mateixos investigadors no es posen d'acord ni en l'ordre d'aparició d'aquests esdeveniments ni on comença i acaba cadascun d'ells.

Així doncs, podem considerar que el concepte primari en la definició d'emoció, tant en teories populars com científiques, té a veure amb els sentiments o sensacions subjectives que experimenten les persones -al cap i a la fi, la paraula emoció prové de la paraula llatina que significa *mouere o remouere* . Ara bé, de quantes emocions estem parlant? Com es relacionen entre elles? com es poden mesurar?

2.2.- Principals formes d'emoció

Identificar les formes en què s'experimenten i expressen les emocions ajuda tant a científics com a les persones en general a poder conceptualitzar aquestes i diferenciar-les una de l'altra. Dins del món de la recerca, i basant-nos en la literatura de la que disposem, els enfoc més destacats des dels quals es classifiquen les emocions són: el mètode categòric, el dimensional i el de prototip. A més d'aquests, considerarem l'enfoc dels afectes de vitalitat (veieu Juslin i Sloboda 1998) el qual és considerat de gran utilitat en l'intent de donar una explicació a les respostes emocionals dels oients a partir de la música.

D'acord amb *l'enfoc categòric*, les persones experimenten emocions com a categories que són diferents una de les altres. Essencial per aquest enfoc és el concepte d'emocions bàsiques; és a dir, la idea de que hi ha un número limitat de categories d'emocions innates i universals, de les quals deriven tots els altres estats emocionals (veieu Ekman 1992; Izard 1977; Oatley 1992; Plutchik 1994; Power & Dalgleish 1997; Tomkins 1962). Les emocions bàsiques són considerades com adaptacions al medi per a fer front a emergències de la vida en un temps de resposta curt: "quan no hi ha una solució totalment racional per a un problema o acció, les funcions de les emocions bàsiques ens porten en una direcció millor que una elecció a l'atzar" (Johnson Laird & Oatley 1992, p. 201). Les emocions bàsiques solucionen problemes amb velocitat més que amb precisió. Aquestes poden ser vistes com "algorismes fugaços i ràpids" que fan front a condicions en un temps, coneixement i capacitats cognitives limitades.

S'han suggerit molts criteris que permetin distingir les emocions bàsiques respecte a les emocions "secundàries" o "complexes", dels quals els més destacats són: (a) tenen funcions diverses que contribueixen a la supervivència individual. (b) es troben en totes les cultures, (c) són experimentades com estats d'ànim únics, (d) apareixen aviat en el desenvolupament, (e) s'associen a diversos patrons de canvis psicològics (f) es poden donar en altres primats, i (g) tenen diverses expressions emocionals.

Es considera les emocions "secundàries", en canvi, com a emocions fundades a partir d'emocions bàsiques, però que impliquen "barreges" d'aquestes emocions bàsiques (Plutchik 1994), o avaluacions cognitives que succeeixen a la vegada amb emocions bàsiques (Oatley 1992).

El concepte d'emocions bàsiques ha estat criticat per diversos motius, sobretot perquè diferents investigadors han escollit altres conjunts d'emocions bàsiques (Ortony & Turner 1990). No obstant, tal i com diu Oatley (1992), "el problema no té a veure amb votacions, sinó amb teoria. Tot depèn de com definim les emocions (p. 104)".

Per tant, si un presta atenció als investigadors que consideren les emocions en termes de la seva importància funcional, trobarem un acord raonable entre ells pel que fa almenys a cinc emocions bàsiques: felicitat, tristesa, enuig, por i fàstic (Kemper 1987; Plutchik 1994; Power & Daigleish 1997).

L'enfoc dimensional es centra principalment en les característiques que distingeixen unes emocions de les altres, identificant aquestes en funció de la seva col·locació dins d'un espai bidimensional o, en alguns casos, tridimensional. Les dimensions més

utilitzades dins d'aquest mètode són valència i activitat, en el cas d'un espai bidimensional, o valència, activitat i potència en el cas del tridimensional. La recerca d'aquestes estructures dimensionals es remunta a Spencer (1890; citat a Izard 1977), però les formulacions més influents d'aquest concepte van ser proporcionades per Wundt (1897), Woodworth (1938) i Schlosberg (1941). Les fonts més comuns de les estructures dimensionals són judicis de semblança de paraules emocionals o expressions facials, que s'analitzen mitjançant anàlisi factorial o escalat multidimensional.

El model circumplex de Russell (1980) ha generat la majoria d'investigacions des d'aquest enfoc. Aquest model consisteix en una estructura circular bidimensional en la qual s'usen les dimensions d'activitat i valència.

Dins aquesta estructura, les emocions es relacionen inversament. El model circumplex capta dos aspectes importants de les emocions: que varien en el seu grau de similitud i que certes emocions (content, trist), sovint són considerades com bipolars. Aquesta mateixa estructura es pot trobar en molts àmbits de recerca, incloent la música (Thayer, 1986), amb el que es dona a entendre que aquest model capta algunes característiques fonamentals de les respostes emocionals (veure Figura 1).

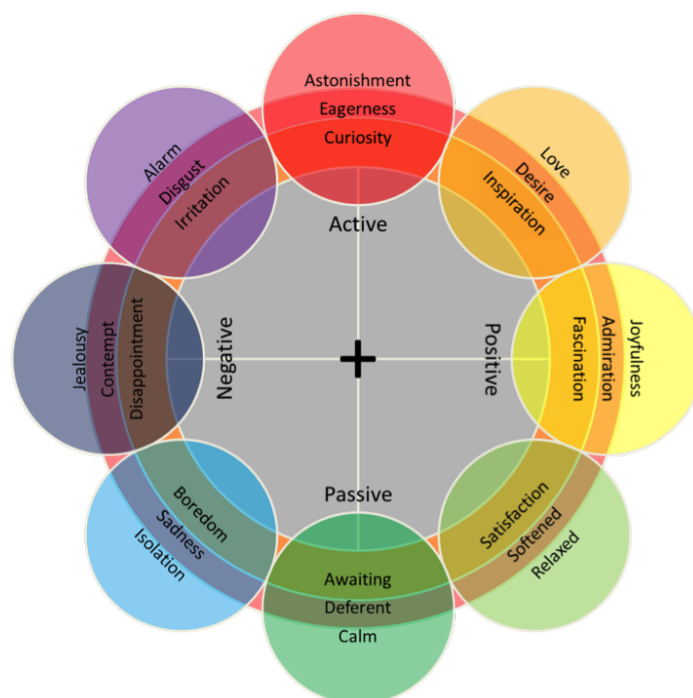


Fig.1 Model circumplex de Russell per a la classificació d'emocions;. En l'eix horitzontal es situa la variable valència, i en l'eix vertical l'activitat.

Un dels punts forts que presenta el model circumplex és que presenta una forma simple i eficaç d'organitzar les emocions en termes de les seves avaluacions afectives (agradable o desagradable) i reaccions psicològiques (alta o baixa excitació). D'una altra banda, els enfoc dimensional han estat criticats per considerar que desdibuixen importants distincions psicològiques, i per tant enfosqueix aspectes importants del procés emocional (Lazarus 1991). En el model circumplex de Russell, per exemple, les emocions que estan situades en la mateixa posició en la matriu circular poden, en realitat, ser diferents. Així, enuig i por són dues emocions que comparteixen uns graus d'activitat i valència similars, però en realitat són molt diferents en quant a les conseqüències per l'organisme. Molts investigadors consideren els enfoc categòrics i dimensionals com a complementaris entre sí (Nyklicek 1997); a més, els models dimensionals poden ser especialment útils per a capturar els continus canvis en l'expressió emocional que succeeixen durant l'escolta d'una peça musical.

El *mètode de prototip* es basa en el treball de Rosch (1978). La idea d'aquest enfoc és que el llenguatge i les estructures de coneixement associades al llenguatge, donen forma a com les persones conceptualitzen o categoritzen la informació. Aquest principi bàsic es pot aplicar a les emocions, de forma que aquestes es puguin categoritzar en diferents nivells en funció de la pertinença a una categoria en particular. Un exemple d'estructura prototípica es pot observar a la Figura 2:

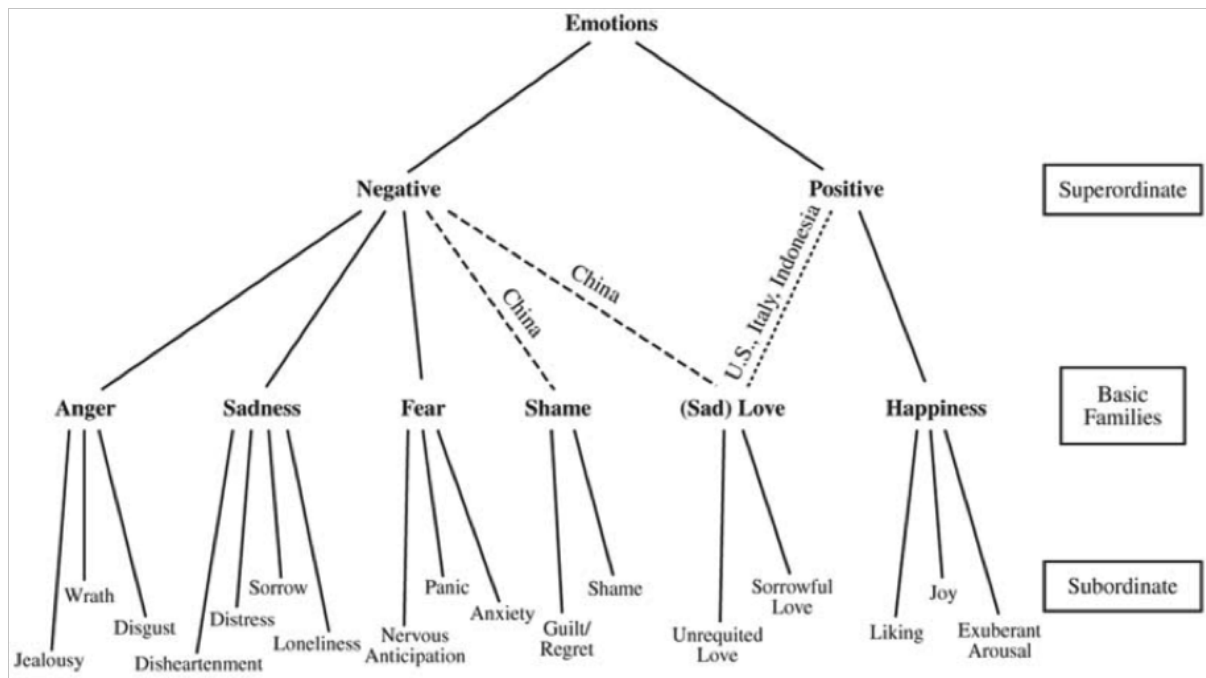


Fig. 2. Categorització de les emocions segons el model de prototip, a "Cross-Cultural Similarities and Differences in Emotion and Its Representation: A Prototype Approach", per R. Shaver, S. Wu i J.C Schwantz, 1992, Newbury Park, CA: Sage.

La dimensió vertical de l'estructura mostra les relacions jeràrquiques entre categories d'emocions. El nivell més generalitzador és el nivell en l'ordre superior, que es defineix segons la valència (positiva o negativa) de les emocions dins una categoria particular. El nivell mitjà representa els prototips o categories d'emocions de nivell bàsic, que engloben les representacions mentals de totes les nostres emocions dins una categoria concreta. El nivell en l'ordre inferior consisteix en totes les altres emocions relacionades amb un prototip en particular. Finalment, la dimensió horitzontal mostra les relacions entre membres de la mateixa categoria *prototip*.

Els crítics del mètode de prototip argumenten que el coneixement de les persones envers les emocions és insuficient per a poder capturar una estructura de les emocions tant definida i jerarquitzada, a la vegada que entenen que no hi ha suficientment acord envers quines emocions es poden considerar prototips de nivell bàsic, de forma que es pot donar el cas que certes emocions creuin fronteres en l'estructura de prototip. No obstant, els partidaris d'aquest enfoc accepten que les fronteres entre categories

d'emocions puguin ser a vegades *borroses*, argumentant que les emocions no poden ser definides en termes d'un conjunt de condicions suficients i necessàries.

A part dels enfocaments més tradicionals que ja hem definit, hi ha una altra classe de qualitats que, malgrat i ser dubtoses en la seva condició d'emocions, semblen ser particularment rellevants en el cas de la música. Stern (1985) va introduir el concepte d'*afectes de vitalitat* per descriure a un conjunt de qualitats difícils de trobar, relacionades amb la intensitat, forma, contorn i moviment. Aquestes característiques es descriuen en termes dinàmics com ara *crescendo*, *fugaç*, *explosiu*, *diminuendo*, etc. Aquestes qualitats no són emocions, sinó "formes" abstractes de sensacions que succeeixen a la vegada o en absència de les emocions adequades. Els afectes de vitalitat són "amodals" en el sentit que són comuns en tots els modes d'expressió. Stern (1985) suggereix que els afectes de vitalitat són de particular importància en els actes comunicatius primerencs de la mare i l'infant. Mare i infant responen un a l'altre ajustant i adaptant contínuament la intensitat, temps i contorn dels seus actes expressius. Aquest procés d'adaptació constant dels esdeveniments gestuals es coneix com *sintonització*. Tot i que el concepte d'afectes de vitalitat és imprecís, sembla capturar aspectes importants de la expressió musical.

Fins ara hem pogut conèixer les formes bàsiques de manifestació de les emocions, les seves possibles categoritzacions i els principals mecanismes implicats en el nostre dia a dia però, quina relació tenen aquestes amb la música? s'evidencien les mateixes emocions i es veuen implicats els mateixos mecanismes en l'escolta d'una obra musical? existeix alguna diferència entre les emocions que expressa la música i el que es sent amb la música (inducció)? En els següents apartats intentarem donar resposta a partir d'un recull de la *recent* literatura existent en aquest àmbit.

2.3.- Com podem mesurar les emocions?

L'emoció és tant un concepte del dia a dia com un constructe científic. Com a tal, implica un conjunt de coneixements implícits i explícits (Plutchik 1980). El coneixement implícit s'engloba en les anomenades *teories populars* de l'emoció. Tothom sap alguna

cosa de les emocions. Pensem que són poderoses forces que afecten el nostre comportament i pensaments intensament, que algunes emocions senten bé i altres no tant, i que hi ha persones més *emocionals* que altres. El coneixement explícit té a veure amb estudis de problemes com ara de quina forma es desenvolupen les emocions, si les emocions estan associades amb canvis psicològics, o com jutgem les emocions en funció de les expressions a nivell facial o vocal. L'emoció com a constructe científic es dedueix a partir de tres tipus de proves: (a) *autoinformes*; (b) *conducta expressiva* i (c) *mesures fisiològiques*.

La forma més comuna –i aparentment senzilla – de mesurar respostes emocionals en humans adults és a través dels *autoinformes*: llistes de control adjectiu, escales de qualificació, questionaris o descripcions lliures. Aquest enfoc està associat a nombrosos problemes, com ara la relació imperfecta entre les emocions i les paraules que les defineixen, i el problema relacionat amb la tria de paraules a incloure en les llistes de control adjectiu, escales o questionaris. Tanmateix, els autoinformes són probablement les proves més directes, i són la solució a certs aspectes de la vida emocional als quals no es pot arribar d'una altra forma que no sigui aquesta.

Donat que l'ús d'autoinformes no és sempre possible – o fiable -, una alternativa a aquest pot ser la mesura de les diferents formes o productes de la conducta expressiva. Aquesta mesura es pot realitzar mitjançant els dibuixos (Smith & Williams 1999), o l'anàlisi dels gestos facials, la vocalització i el llenguatge del cos. Probablement, el cas més famós el trobem en els estudis interculturals d'Ekman (1973) envers l'expressió facial de les emocions. Un problema que ens podem trobar en realitzar aquesta prova és que les emocions no sempre van acompanyades de conductes expressives. Tanmateix, les conductes expressives sovint es produeixen en absència d'emocions, ja que a vegades la gent fa servir aquest tipus de comportaments intencionadament per comunicar certa informació a altres persones.

El tercer tipus d'estudis usats per extraure informació envers les emocions implica la mesura fisiològica de les emocions. Aquest tipus reflecteix en part el treball de William James, qui va afirmar que l'emoció es basava en la percepció dels canvis corporals interns. Des de llavors, els investigadors han desenvolupat una sèrie de d'índexs fisiològics per mesurar les emocions, incloent la freqüència cardíaca, la respiració, la conductància de la pell, la tensió muscular, l'electrocardiograma, la pressió sanguínea i l'electroencefalograma (per revisió, veieu Cacioppo, 1993). Hi ha diversos problemes

amb les mesures fisiològiques, dels quals el més important probablement sigui que els canvis autonòmics sovint es produeixen en absència d'emocions (degut principalment a que cada sistema fisiològic té moltes funcions diferents dins del cos); per tant, és difícil establir relacions clares entre estats emocionals i respostes fisiològiques.

2.4.- Quina relació té la música i les emocions?

Si fem un repàs a la història de la música, segur que trobem multitud d'exemples en els quals s'hagin establert relacions entre la música i les emocions; ja en la Grècia Antiga, es considerava que els modes musicals tenien un efecte sobre les emocions (una creença que es va seguir mantenint en el període medieval, en la classificació de modes segons l'emoció evocada creada per Guido d'Arezzo).

Malgrat la multitud d'esments que s'han produït en referència a aquesta relació al llarg de la història pocs tractats ho han plasmat en els seus redactats (com a excepcions, veieu Fridja 1989; Lazarus 1991; Oatley 1992; Zajonc 1994). Això pot reflectir el supòsit que les emocions experimentades mentre s'escolta música són en cert sentit diferents d'altres emocions. En efecte, hi ha diferències importants entre les emocions musicals i altres emocions, tant en els antecedents com en les conseqüències, encara que en sí mateixes no impliquen que les emocions pròpiament siguin diferents. Una altra raó per a l'abandonament de la música en la literatura de les emocions podria ser que tals emocions es consideren d'alguna forma menys importants –potser perquè els científics no s'han adonat, fins dates recents, que la música compleix funcions vitals en les vides de moltes persones (Sloboda & O'Neill, 1997).

Les emocions estètiques –reaccions emocionals a pel·lícules, drames, música o pintura- mai han estat el principal cavall de batalla de la psicologia (Lazarus 1991). Tot i que els psicòlegs han reconegut que l'art, malgrat la seva naturalesa il·lusòria, pot evocar fortes reaccions emocionals, solament alguns han considerat aquesta relació digna d'estudi. Fins i tot en la psicologia de la música contemporània, l'estudi de les emocions s'ha mantingut en un segon pla (Gabrielsson 1993).

No obstant, en la última dècada s'ha pogut observar un augment de l'interès per aquest àmbit d'estudi. Es creu que aquest interès pot ser degut a que la psicologia musical, per poder progressar, ha pres una major consciència de la necessitat d'enfrontar aquest problema tenint en compte la forma en que les persones experiencien la música (Juslin 1997).

Les emocions musicals són, a primera vista, de dos tipus: d'una banda, aquelles emocions que es refereixen al valor estètic de la música. Els psicòlegs tenen, per desgràcia, poc a dir sobre l'estètica de la música, mentre els filòsofs han estat discutint sobre aquest tema durant segles (Budd 1985). Es considera que una resposta estètica és una experiència personal intensa que dona una idea de la naturalesa de la vida i que implica components emocionals, cognitius i socials (Konecni 1979). En el camp de la "nova" estètica empírica, no obstant, els investigadors s'han centrat en relacions més mundanes, com el gust, la preferència i les propietats que donen informació d'estímul com són la complexitat, la familiaritat o expectació (Berlyne 1971; North & Hargreaves 1997; per a una revisió de les preferències musicals, veieu Russell 1997). D'una altra banda trobem aquelles emocions que s'indueixen o expressen mitjançant la música però que es desvinculen del valor estètic d'aquesta. Aquestes són les emocions en les que ens centrarem; no obstant, els dos tipus d'emocions no són del tot independents entre elles, i es creu que per entendre els aspectes psicològics de la música i les emocions, és necessari que els investigadors realitzin i mantinguin aquesta connexió de forma explícita (Juslin 1997).

L'estudi de les emocions en relació amb la música pot ser problemàtica en certs punts. Primerament, les reaccions emocionals s'entenen habitualment en termes de les seves funcions adaptatives en relació amb la supervivència biològica. Ja que les reaccions musicals no estan directament lligades amb aquesta supervivència biològica, sembla difícil explicar-les i definir-les des del punt de vista de les teories de les emocions de les quals tenim coneixement (Dowling & Harwood 1986). A més, els científics que volen estudiar les reaccions emocionals envers la música es troben amb el problema de la gran variabilitat entre individus, i a través del temps en els individus (Sloboda 1996). Finalment, els experiments que intenten mesurar les respostes afectives a partir de la música poden impactar tant en el procés d'escolta que destrueix l'objectiu de mesura real –el problema de la *reactivitat* (Neale & Liebert 1986).

No obstant això, alguns autors suggereixen que la música és un excel·lent medi per estudiar les emocions. Per exemple, Gaver i Mandler (1987) argumenten que fer ús de l'estudi dels estímuls musicals és ecològicament vàlid, ja que les persones estan acostumades a fer judicis sobre la música i les seves respostes afectives. Tanmateix, consideren que la música té una estructura rica però força coneguda, el que facilita i ajuda a entendre les reaccions musicals que produeix o transmet. Un altre avantatge de l'ús de la música com a estímul és que es desentèn de moltes consideracions ètiques associades normalment a altres tipus d'estímuls.

Dels mecanismes psicològics que estan involucrats en la inducció d'emocions en la música, Juslin considera que "els sis mecanismes més destacats són: (1) *reflexos del tronc encefàlic*, (2) condicionament avaluatiu, (3) *contagi emocional*, (4) *imatges visuals*, (5) *memòria episòdica* i (6) *l'expectació musical*". Es creu que aquests mecanismes, juntament amb l'avaluació cognitiva, poden explicar la majoria d'emocions induïdes per la música en la vida quotidiana (Juslin, 2008). Aquestes són les característiques més destacades de cadascun dels mecanismes esmentats:

- *Reflexos del tronc encefàlic*. Es refereix al procés en el qual el tronc encefàlic pren característiques acústiques fonamentals de la música per a induïr emocions, donant lloc a missatges potencialment importants. En definitiva, aquests missatges reflecteixen l'impacte de les sensacions auditives –música com a so en el seu sentit més bàsic i simple.

- *Condicionament avaluatiu*. Té a veure amb el procés mitjançant el qual una emoció és induïda en el moment en què els estímuls s'emparellen amb altres estímuls positius o negatius de forma repetida. Per exemple, una peça musical en particular pot coincidir repetidament en moments en què estàs feliç; això pot fer que en un futur aquesta música et transmeti felicitat.

- *Contagi emocional*. En aquest cas, una emoció és induïda a partir d'una peça musical degut a que l'oïent percep l'expressió emocional d'aquella música, imitant l'expressió internament mitjançant qualsevol mecanisme de retroalimentació perifèrica dels músculs o l'activació directa dels elements emocionals rellevants en el cervell. Per

exemple, la música podria reflectir una expressió trista (tempo lent, to baix, volum baix), que indueixi tristesa en l'oient.

- *Imatges visuals*. Es refereix al procés en el qual l'escolta musical evoca en l'oient imatges visuals (per exemple, un camp verd meravellós). Les emocions experimentades seran el resultat de la interacció entre música i imatge.

- *Memòria episòdica*. Té a veure amb la inducció d'emocions a partir del record particular d'un event viscut per l'oient en el moment de l'escolta musical.

- *Expectació musical*. En aquest procés, una emoció és induïda en l'oient quan algun element o característica de la música viola, retarda o confirma el que esperava l'oient escoltar com a continuació de la peça musical. Per exemple, la progressió E-F# estableix una expectació cap al G# (Sloboda 1992). Si això no succeeix, l'oient es pot veure sorprès.

2.5.- Com responen els oients a la música? (percepció vs inducció)

La majoria dels esforços en l'àmbit de la recerca musical es centren, actualment, en entendre com els oients responen a la música, plantejant preguntes com ara si la música expressa i indueix emocions, si aquestes són clarament diferenciables o quins són els mecanismes que fan que es reaccioni emocionalment a aquesta. A partir de la literatura existent en aquest àmbit, intentarem respondre a aquestes preguntes.

Hi ha moltes fonts d'emoció en la música (Sloboda & Juslin, 2001); és a dir, un event musical pot expressar i evocar emotions de diferents formes. Per exemple, algunes emocions es poden despertar a partir de característiques estructurals de la música i en canvi, altres, reflectir associacions personals que no necessàriament despertaran les mateixes emocions. Així doncs, no especificar quina font d'emoció s'està estudiant pot portar a controvèrsies injustificades amb altres fonts. Les fonts emocionals es poden veure afectades i implicades en un gran número de variables causals tant en la música

com en la persona i la situació (Jorgensen, 1988), pel que és de gran importància en la recerca definir quines d'aquestes variables es tindran en compte i quines no.

D'una altra banda, és important fer una distinció entre percepció i inducció d'emocions, entenent com a percepció el detectar una càrrega comunicativa emocional en la música i com a inducció el sentir emocions en resposta a la música. Aquesta distinció, coneguda des dels temps de la Grècia antiga, sovint, però no sempre, es realitza en l'àmbit de la recerca moderna. És important fer aquesta distinció per tres motius. Primer, els mecanismes subjacents poden ser diferents segons els processos involucrats. Segon, mesurar l'emoció induïda és més difícil que mesurar l'emoció percebuda, per tant els investigadors s'han d'adaptar en conseqüència. Tercer, els tipus d'emocions expressades i percebudes de la música poden ser diferents de les emocions induïdes a partir de la música.

La percepció emocional és relativament fàcil de mesurar, i és un procés "cognitiu" en el sentit que es pot dur a terme sense cap implicació emocional per part de l'oient, això fa que no hi hagi una resposta correcta o incorrecta a l'emoció percebuda en l'escolta d'una peça musical. No obstant, als investigadors els interessen els casos en què les emocions puguin ser percebudes de forma similar per diferents oients (o percebuda com el compositor o l'intèrpret vol), de forma que aquesta s'avalua en termes d'(a) *acord amb l'oient* (on es diu que l'emoció es considera fiable si hi ha un cert nivell d'acord entre els oients sobre el que expressa la música) i *precisió* (la qual es refereix al reconeixement "correcte" per part dels oients, d'acord amb algun "criteri independent", tal com la intenció del compositor o intèrpret). La majoria de les investigacions es basen en mesures d'acord amb l'oient, ja que normalment és difícil d'obtenir índexs fiables de les intencions expressives dels compositors.

El resultat obtingut d'un recull de més de 100 estudis (Juslin i Laukka, 2004) mostra una bona precisió en els judicis dels oients a diferents fragments musicals. No obstant, aquests també mostren millors resultats quan se'ls hi demanen categories emocionals en un sentit ampli, i no tant quan es busquen emocions amb característiques més específiques, la qual cosa posa en evidència que la precisió és limitada (Campbell, 1942; Downey, 1897; Juslin 1997c).

L'acord dels oients envers les emocions expressades depèn de molts factors (p. ex, la peça musical, l'estil musical, el format de resposta, el procediment), però s'ha observat

més robustesa en els resultats quan es marginaven les respostes en funció de l'entrenament musical, edat, i sexe de l'oient. En el cas concret de no tenir entrenament musical, es creu que l'habilitat per reconèixer emocions discretes està relacionada amb mesures d'"Intel·ligència Emocional". Com ja hem dit anteriorment, la percepció d'emocions en els oients es porta a terme normalment mitjançant experiments en els quals els oients han de jutjar l'emoció percebuda mitjançant *resposta única*, la *selecció d'adjectius relacionats* amb l'emoció, *qualificació d'adjectius* relacionats amb emocions (en una escala amb un valor mínim i màxim) i *descripció lliure*. Com era d'esperar, la variabilitat en les respostes és més gran en el cas de la descripció lliure que en la resposta única (Juslin, 1997c; Rigg 1937). A més, l'acord dels oients és més gran en unes emocions (p. ex felicitat, tristesa) que en unes altres (p. ex. gelosia), el que indica clarament que la música pot transmetre millor unes emocions que altres.

2.6.- El reconeixement d'emocions en la música

El reconeixement d'emocions en la música ha estat un tema d'estudi en la última dècada.

Des d'un punt de vista de l'enginyeria i la ciència afectiva, un dels principals punts d'interès s'ha centrat en desenvolupar models d'aprenentatge automàtic capaços de reconèixer o categoritzar emocions en la música.

Els models d'aprenentatge automàtic són el resultat d'una sèrie de tècniques computacionals que, aplicades a un arxiu musical, tracten de categoritzar emocions (com per exemple tristesa, enuig, felicitat i relaxació) a partir de l'anàlisi de les seves característiques més rellevants.

El desenvolupament d'un model de reconeixement d'emocions en la música consta de les següents fases:

- *Creació d'una col·lecció d'arxius d'àudio amb música*: Tant important és l'anàlisi i avaluació de les senyals d'àudio com l'existència de col·leccions d'àudio adients a la recerca; així doncs, les col·leccions han de complir els requisits necessaris

per a què les preguntes plantejades a la recerca puguin ser resoltes amb el material d'àudio generat. Els aspectes més importants a tenir en compte en la creació d'una col.lecció d'àudio són: *tamany de la mostra, idioma i contingut emocional*. Si voleu més informació envers les col.leccions d'àudio, veieu la base de dades de col.leccions d'àudio de l'Associació HUMAINE per la Ciència Afectiva (AAAC)¹ la base de de dades de la que es disposa en l'àmbit de la recerca

- *Mesura de les respostes emocionals percebudes a partir de l'escolta dels arxius de la col.lecció*: L'avaluació de les respostes emocionals és un element bàsic en el procés de creació de models d'aprenentatge automàtic; com hem parlat anteriorment (veieu secció 2.3) les principals formes de mesura de les emocions com a constructe científic són (a) *autoinformes*; (b) *conducta expressiva* i (c) *mesures fisiològiques*.

- *Extracció de descriptors per cadascun dels arxius musicals de la col.lecció*: els descriptors o "features" són mesures contínues o discretes realitzades a partir de l'anàlisi de la senyal d'àudio, sigui amb caràcter musical o no, que ajuden a donar una explicació i extreure conclusions envers els esdeveniments musicals o sons existents en l'arxiu sonor o musical.

- *Categorització de les emocions i definició de descriptors rellevants a partir de tècniques d'aprenentatge automàtic*: en aquesta part del procés es recullen les anotacions (ratings) obtingudes en la mesura d'emocions i, mitjançant tècniques d'aprenentatge automàtic, es cerquen possibles relacions amb els descriptors analitzats.

¹ <http://emotion-research.net/association>

2.7.- Els descriptors d'àudio

Un fragment d'àudio pot ser analitzat tant en el domini temporal com en el freqüencial (si apliquem la transformada de Fourier a la senyal).

El processat digital d'una senyal d'àudio per l'anàlisi consta de a) *enfinestrat* i b) *solapament de les finestres*; en l'*enfinestrat*, es divideix la senyal en x parts, a les quals se li aplica una finestra que s'adequi a les necessitats de l'anàlisi en funció de la resolució temporal o freqüencial desitjada, i en el *solapament*, es seleccionen les finestres i es sotmeten a un procés de solapament que redueixi l'efecte dels lòbuls de l'enfinestrat.

Quan una senyal d'àudio s'analitza en aquest nivell diem que l'anàlisi que estem fent és de baix nivell i que per tant els descriptors d'aquest anàlisi també ho són; aquest tipus de descriptors es coneixen amb el nom de *Low-Level Descriptors (LLD)*. Els LLD són un element bàsic en la detecció automàtica d'emocions, ja que permet analitzar la variació en les mesures dels atributs que es veuen afectats per un canvi en l'emoció representada i que, per tant, es consideren útils en la tasca de diferenciació d'emocions.

A continuació es mostra un llistat dels descriptors utilitzats per l'anàlisi de la veu en l'experiment realitzat per Eyben envers les emocions en la música cantada (Eyben et al., 2015):

- Proporció d'energia per sota dels 500 Hz respecte a l'energia total.
- Proporció d'energia per sota dels 1000 Hz respecte a l'energia total.
- Freqüència del primer, segon i tercer formant
- Ample de banda del primer formant
- Alpha ratio (dB) entre l'energia per sota i per damunt dels 1000 Hz.
- Distància entre el primer harmònic i el segon harmònic (H1-H2 LTAS)
- Distància entre el primer harmònic i el tercer harmònic (H1-H3 LTAS)
- Índex Hammerberg (dB), entès com la diferència d'energia màxima entre 0 i 2000 Hz i l'energia màxima entre 2000 i 4000 Hz.
- Pendent espectral (1000 – 5000 Hz)

- Planura espectral
- Centroide espectral
- Freqüència fonamental (f_0)
- Jitter (%), entès com la mitjana de la diferència entre períodes respecte al període mig
- Shimmer (%): entès com la mitjana de la diferència absoluta de les amplituds respecte a l'amplitud mitja.
- HNR: Harmonic-to-noise-ratio, entès com la diferència de l'energia dels harmònics respecte al soroll.
- Autocorrelació mitjana
- Tempo, entès com el nombre de pulsos per minut..

Els LLD ens permeten calcular fàcilment moltes característiques del so, però hi ha altres característiques de la música que poden ajudar a donar informació complementària i necessària per un correcte anàlisi de les emocions. Aquestes són *el tempo, el mode, l'harmonia, la tonalitat, el to, la micro-entonació, el contorn, la intervàlica, el ritme, la intensitat sonora, el timbre, la sincronia, l'articulació els accents en notes específiques, els atacs i caigudes i el vibrato*. Es considera que aquestes característiques, configurades i combinades de la forma adequada, poden ajudar a la categorització d'emocions expressades en la música (Juslin, 2001), però encara no tenim tècniques robustes per a determinar-les a partir de l'àudio.

Juslin, a partir de lectura existent proposa una aproximació de la configuració d'aquestes característiques en funció de l'emoció bàsica expressada. Veieu a continuació:

- *Felicitat*: tempo ràpid, poca variabilitat del tempo, mode major, harmonia consonant i simple, nivell de so mig-alt, poca variabilitat de nivell de so, to alt, molta variabilitat del to, to ascendent, intervals de 4^a i 5^a, augment de la micro-entonació, augment dels formants del cantant, articulació amb staccato, gran variabilitat de l'articulació, ritmes fluïts i suaus, timbre brillant, atacs ràpids, variabilitat de la sincronia petita, contrastos aguts entre notes llargues i curtes, ratio del vibrato mig-ràpid, extensió del vibrato mitjana, regularitat microestructural

- *Tristesia*: tempo lent, mode menor, dissonàncies, nivell de so baix, variabilitat del nivell de so moderada, to baix, rang de to petit, to descendent, entonació plana o caient, intervals petits (segones menors), formants del cantant baixes, articulació amb legato, poca variabilitat en l'articulació timbre apagat, atacs lents, variabilitat sincronia (rubato), pocs contrastos entre notes llargues i curtes, pauses, vibrato lent, poca extensió del vibrato, ritardando, irregularitat micro-estructural.

- *Enuig*: Tempo ràpid, poca variabilitat del tempo, mode menor, dissonància, nivell de so alt, poca variabilitat del nivell de so, to alt, poca variabilitat del to, to ascendent, intervals de 7ena i 4^a, augment dels formants del cantant, articulació stacatto, variabilitat moderada de l'articulació, ritme complex, canvis rítmics sobtats (síncopes), timbre agut, soroll espectral, atacs i caigudes ràpides, poca variabilitat de la sincronia, accents en notes inestables a nivell tonal, forts contrastos entre notes llargues i curtes, accelerando, vibrato mig-ràpid, gran extensió del vibrato, irregularitat micro-estructural.

- *Por*: Tempo ràpid, variabilitat del tempo gran, mode menor, dissonància, nivell de so baix, variabilitat del nivell de so gran, canvis ràpids en el nivell de so, to alt, to ascendent, rang del to ample, contrastos grans de to, articulació stacatto, variabilitat de l'articulació gran, ritmes bruscos, timbre suau, variabilitat de la sincronia molt gran, pauses, atacs del to suaus, ràtio del vibrato alta, extensió del vibrato petita, irregularitat micro-estructural.

- *Tendresa*: Tempo lent, mode major, consonància, nivell de so mig-baix, variabilitat del nivell de so petita, to baix, rang de to força petit, formants del cantant baixos, articulació legato, variabilitat de l'articulació petita, atacs del to lents, timbre suau, variabilitat de la sincronia moderada, contrastos suaus entre notes llargues i curtes, accents en notes tonalment estables, vibrato mig-ràpid, extensió del vibrato petita, regularitat micro-estructural.

2.8.- Les tècniques d'aprenentatge automàtic

Les tècniques d'aprenentatge automàtic a utilitzar en una recerca poden variar en funció de l'objectiu d'aquesta. Algunes de les tècniques més utilitzades són:

- *Decision Trees*: Els arbres de decisió són una eina de suport que categoritza les decisions i les seves possibles conseqüències i les representa en gràfics en forma d'arbre. Veieu la figura 3:

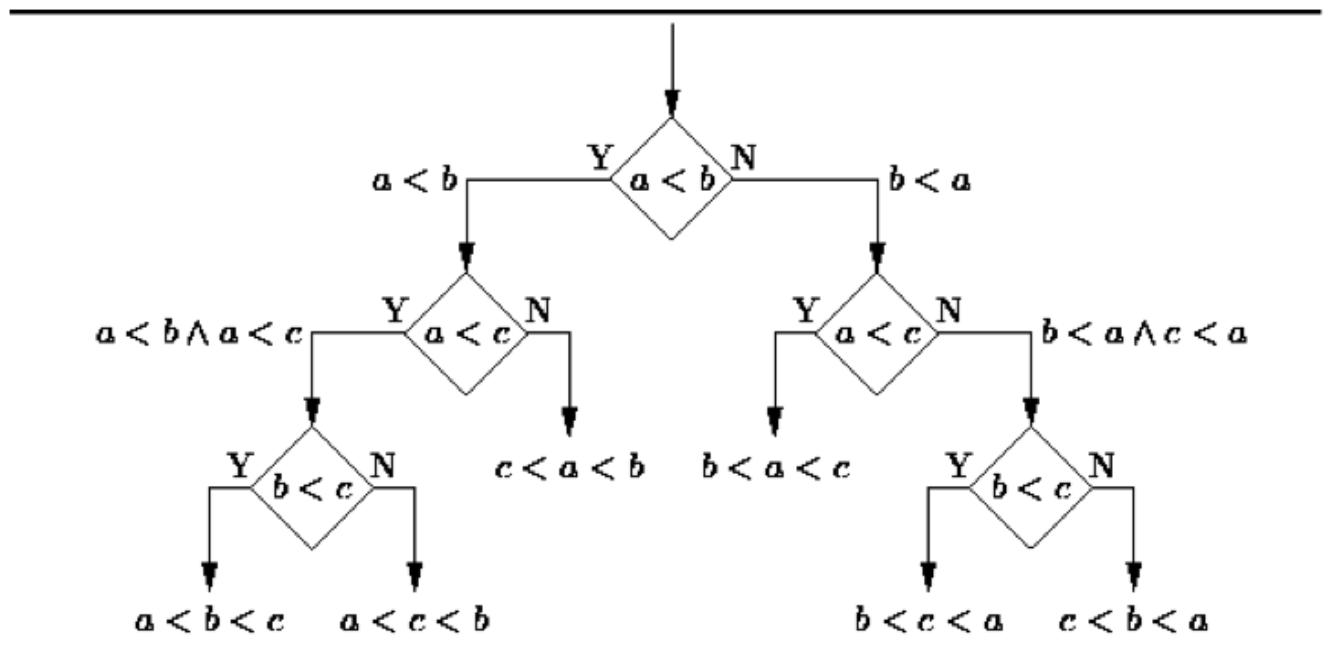


Fig3. Arbre de decisions

- *Naïve Bayes Classification*: Basats en el teorema de Bayes, aquests classificadors probabilístics simples cerca relacions entre característiques a partir d'una variable concreta (*naïve*).

- *Ordinary Least Squares Regression*: També conegut com regressió lineal, aquest mètode cerca predir o ajustar una sèrie d'observacions a un model conceptualment simple com pot ser una línia recta; hi han múltiples estratègies per fer-ho, i una d'elles és la *suma de quadrats*.

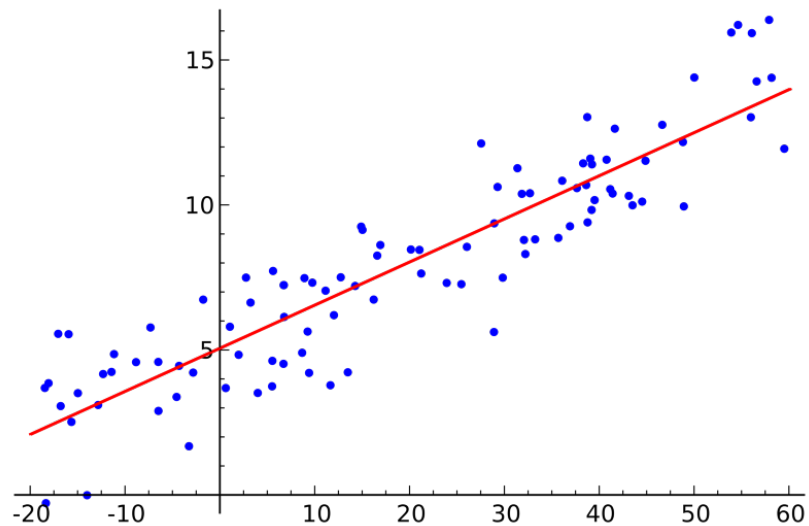


Fig. 4. Ordinary Least Square Regression

- *Logistic Regression*: La regressió logística és una forma estadística potent en el modelatge d'un resultat a partir d'una o més variables explicatives. Aquesta mesura la relació entre la variable dependent categòrica i una o més variables independents.

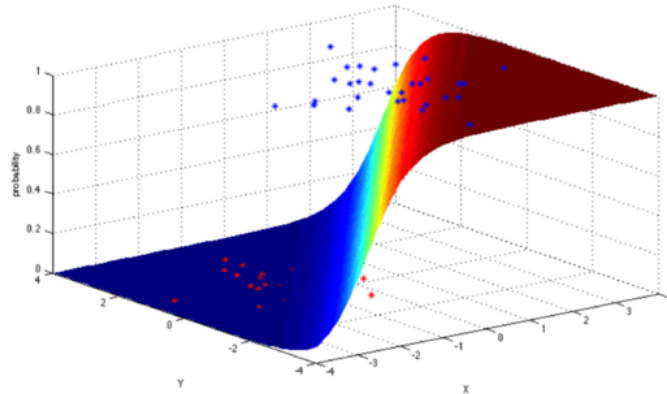


Fig. 5. Logistic Regression

En general, les regressions es poden aplicar en molts processos del món real, com poden ser la mesura del grau d'èxit de campanyes de màrqueting, la predicció dels ingressos d'un producte o fins i tot si hi haurà un terratrèmol en un dia particular.

Un dels valors més importants en els mètodes de regressió és el coeficient de correlació. S'entèn com a correlació a la relació al grau de relació entre les variables explicatives o independents i la variable dependent o variable a predir. El coeficient de correlació comprèn valors entre 0 (mínim grau de correlació) i 1 (o -1) (màxim grau de correlació). A la figura 6 es poden observar exemples de correlació aproximats:

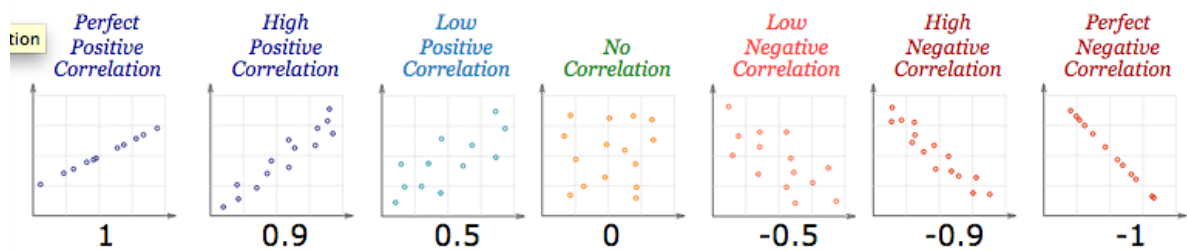


Fig. 6. Aproximació visual dels coeficients de correlació

- *Support Vector Machines*: SVM és un algorisme de classificació binari. A partir d'una sèrie de punts de 2 tipus en un espai dimensional N, l'SVM genera un hiperpla dimensional (N-1) per a separar aquests punts en 2 grups.

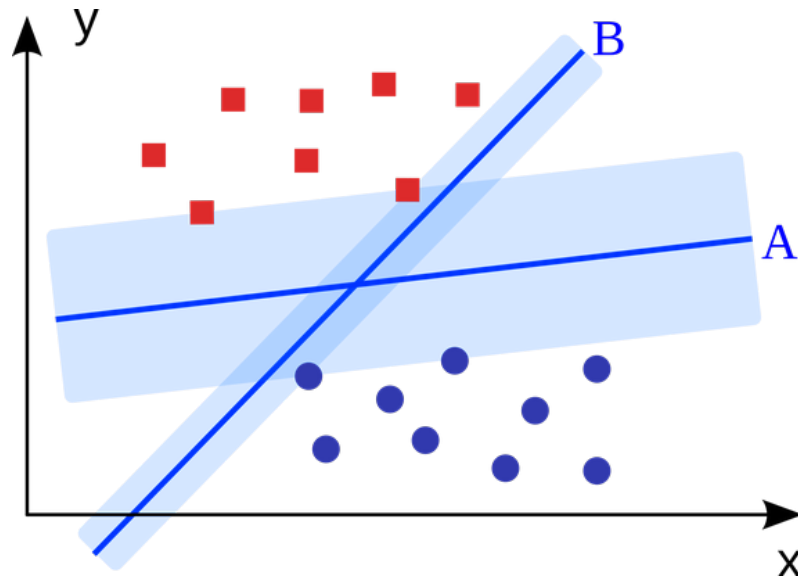


Fig. 7. *Support Vector Machines*

- *Ensemble Methods*: es refereix a combinacions de diferents mecanismes o models d'aprenentatge tant per la creació de classificadors combinats com per la constitució d'ensembles experts basats tots en el mateix classificador, on cada expert es dedica a classificar booleanament una categoria entre les diverses possibles.

2.9.- La producció musical com a creadora d'emocions

La producció musical ha estat l'últim element que ha entrat amb força com a portadora d'emocions; si realitzem una comparativa entre les produccions dels primers enregistraments i les produccions en l'actualitat podrem observar una sèrie de diferències que han estat causa i conseqüència de la forma en que les persones escolten i responen a la música.

Per poder entendre aquest fenomen ens hem de remuntar a abans de l'aparició dels primers enregistraments; si pensem en com la música s'escoltava abans dels primers enregistraments, ens imaginem un context en que el músic interpreta la música i l'oient escolta en directe la interpretació amb una referència espacial relacionada amb el lloc d'interpretació evidentment. La primera evolució en la forma d'escoltar la música sorgeix a partir dels primers enregistraments, en els que la funció principal de la producció era la de permetre que el material gravat es mantingués en el temps més enllà del moment de la interpretació musical; aquest avenç produeix el primer canvi en la forma d'escoltar la música de l'oient, ja que aquest fet fa que la música desvinculi el lloc d'escolta amb el lloc d'enregistrament. La següent evolució en l'àmbit de la producció es produeix amb l'entrada d'elements per al processament del so; en aquesta fase de l'evolució es desenvolupen equipaments que permeten donar al so un caràcter tímbric diferent a l'original en el moment de l'enregistrament, amb el que ja no necessàriament l'oient escolta el que en el moment de l'enregistrament va ser reproduït. Finalment, l'evolució de les tècniques d'enregistrament amplien les eines de les que disposa el productor en els seus treballs, i li donen la posició creativa de la que no disposava fins aquell moment; aquesta posició creativa incrementa les possibilitats de manipulació de la música i la desvinculació de l'oient amb l'espai on es realitza l'enregistrament, amb el que aquest es va adaptant al nou panorama fins al punt de no necessitar aquesta vinculació o connexió visual imaginària de l'espai d'interpretació o enregistrament.

Amb aquest nou canvi de panorama, en el que el productor és considerat el sisè músic i assoleix un rol més creatiu, les possibilitats de producció s'expressen al màxim amb l'objectiu de donar a l'oient una experiència més enllà de la corresponent a la interpretació musical; l'ús de la reverb com a element de recreació ambiental i espacial, l'autotune com a eina per a la correcció del contorn melòdic per a un objectiu

emocional determinat, o l'aplicació de procediments en el processat de la veu per reduir o incrementar el breathiness en són uns clars exemples.

Les investigacions reitzades en aquest àmbit mostren un efecte rellevant dels elements de la producció sobre la percepció de les emocions; en aquest sentit, Fletcher (Fletcher, 2011) en la seva recerca envers el tractament espacial en la percepció emocional de l'oient i Ronan (Ronan, 2016) en la recerca de relacions entre la qualitat de la producció musical i les emocions ho corroboren.

En el panorama musical actual podem trobar exemples d'aquest ús de la producció com a element rellevant en la percepció d'emocions, com ara el tractament espacial en les últimes produccions del grup nord-americà *Band of Horses*, o l'ús de vocoders i harmonitzadors en les últimes produccions del també nord-americà *Bon Iver*.

3.- Objectius

Els objectius principals d'aquest treball són: a) descriure les característiques i problemàtiques més destacables en la classificació d'emocions expressades en la música a partir de la literatura de la que es disposa actualment; b) oferir nova informació en l'àmbit de la recerca musical, basada en l'anàlisi i extracció de característiques musicals i descriptors a partir de la conveniència o no de la separació d'elements d'un fragment musical; i c) obtenir recursos i coneixements envers els procediments a seguir tant en el desenvolupament d'un treball científic, com en l'anàlisi de fragments d'àudio a partir de l'ús i extracció de descriptors i característiques musicals en el reconeixement automàtic d'emocions

4.- Metodologia

En la metodologia s'han utilitzat els procediments i mètodes que puguin donar resposta a les preguntes plantejades a partir dels objectius proposats; així doncs, les tècniques metodològiques han de respondre a preguntes com quines emocions es perceben en els fragments d'àudio, quines emocions varien en funció dels components de la cançó utilitzats o si existeix una sinèrgia dels diferents components d'una cançó de forma que quan es combinen, l'emoció es potencia o per contra es debilita.

4.1.- Creació de la col.lecció d'àudios

La col.lecció s'ha construït a partir d'una sèrie d'arxius multipista obtinguts de la base de dades de la pàgina web Cambridge MT² (Moffat et al., 2015; Witten et al, 2005; Cartwright et al, 2014, Wilson et al., 2016), la qual disposa d'un recull ampli d'arxius de so per a la pràctica en usos varis del món de la tecnologia del so; la majoria d'enregistraments de la base de dades estan realitzats amb disposició microfònica propera i amb els elements de la mescla aïllats, així que és possible dividir els diferents elements d'una cançó en funció de les necessitats.

De cadascun dels arxius multipista s'han extret 4 versions diferents: *veu*, *instrumental*, *veu + instrumental*, *mescla més postproducció*, per a què en els processos posteriors es pugui fer la comparativa de resultats entre les diferents versions.

Aquests arxius multipista, amb les seves diferents versions, han de complir una sèrie de característiques, ja que sense aquestes la recerca podria veure's *adulterada* i no complir els mínims per a donar-li la validesa suficient al procediment:

- *Eliminació del "leakage"*. S'entèn per *leakage* com als elements sonors inclosos en la captació microfònica que estan fora de l'objectiu a enregistrar. (p.ex: el so del piano en la captació microfònica del contrabaix en un trio de jazz). Així doncs, els arxius multipista, en el moment de, per exemple, aïllar la veu, han d'assegurar que el micròfon que captava la veu no es vegi afectat per altres instruments; en cas

² <http://www.cambridge-mt.com>

contrari, la comparativa entre versions no seria fiable.

- *Duració dels arxius similar.* S'ha de limitar la duració dels arxius perquè aquesta no sigui motiu d'esbiaix de la percepció emocional d'aquests. La duració aproximada de cadascun dels arxius inclosos en la col.lecció és de 30 segons.

- *Arxius homogenis a nivell musical:* Cadascun dels arxius ha de mantenir una certa homogeneïtat textural, harmònica, rítmica i d'intensitat sonora, per tal d'assegurar una percepció única.

- *Mateix format d'àudio.* Per assegurar que la codificació perceptual no afecti a la forma en que els oients poden percebre les diferents característiques sonores dels arxius, han de poder escoltar tots els arxius en el mateix format.

- *Major qualitat possible dels arxius.* Els formats de compressió amb pèrdua, segons la seva qualitat, poden afectar a les característiques del so i inclús incloure artefactes que en l'arxiu original amb major qualitat no existien, per això hem d'assegurar la major qualitat; en aquest cas s'ha hagut de buscar un compromís entre qualitat i capacitat de l'arxiu, ja que cadascun d'aquests serà penjat en un servidor per al procés d'avaluació i extracció d'annotadors. Així, els arxius han estat sotmesos a compressió amb pèrdua MP3 però assegurant una qualitat mínima de 192 Kbps.

- *Mateix nivell de sonoritat:* El sonoritat es defineix com el nivell d'intensitat amb la que l'oient percep un so; aquest és un element molt important en la percepció d'emocions, ja que si els nivells de sonoritat són diferents, l'equilibri entre freqüències greus, mitjes i agudes es pot veure afectat, i per tant modificar la percepció emocional d'un arxiu sonor. Així doncs, cadascun dels arxius han estat normalitzats seguint la normativa EBU R128, en la que es requereix una mesura global de sonoritat de -23 LUFS.

- *Representació d'un ventall suficient d'estils i emocions.* Per a que la recerca sigui més rica i generalitzada, és necessari que la col.lecció mostri un ventall d'emocions i estils ampli. Així doncs, els arxius representen estils generals com el jazz, el rock, la música electrònica, el hip-hop o el pop, a la vegada que es representen

diferents emocions en cadascun d'ells, seguint criteris de valència i activitat (tal i com hem explicat a la secció *principals formes d'emoció* de l'estat de la qüestió)

4.2.- Avaluació d'emocions i extracció d'annotadors a partir de la col.lecció

L'avaluació d'emocions i extracció d'anotacions és una part molt important dins la recerca. En aquesta part del procés, es realitza una sèrie de proves a individus aliens a la recerca, als quals se'ls hi demana respondre a una sèrie de preguntes envers l'escolta musical dels arxius sonors de la col.lecció.

El procediment escollit per la mesura de les respostes emocionals és l'auto-informe, a través d'un questionari amb escales de qualificació i preguntes de resposta única o múltiple.

Com ja hem dit anteriorment, una de les dificultats d'aquest procediment és la dificultat per aconseguir la relació perfecta entre les emocions i les paraules o adjectius associats a cadascuna d'aquestes emocions. No obstant, els autoinformes són probablement les proves més directes, i permeten enfocar la mesura de les respostes emocionals a través del que la música expressa, i no les emocions que aquesta indueix en els oients.

La creació del qüestionari s'ha realitzat a partir de la plataforma web PsyToolkit³, la qual permet, a través de funcions preestablertes, dissenyar scripts per a proves científiques amb un gran ventall d'opcions de manipulació d'aquestes. Així, permet incloure arxius audiovisuals, realitzar preguntes amb múltiples opcions de resposta (resposta única, múltiple, *valors* en un interval numèric determinat, descripció lliure) i ordres aleatoris, fer selecció de preguntes per blocs, i altres opcions que poden interessar en un projecte de recerca científica; a més permet la difusió del qüestionari mitjançant un enllaç únic i la recepció d'un feedback de les respostes mitjançant arxius en format .csv.

³ <http://www.psytoolkit.org>

El qüestionari consta de dues parts ben diferenciades: la part que avalua les respostes emocionals en funció de l'escolta musical i una segona part complementària que pretèn relacionar les respostes de la primera part amb acotacions en funció de l'edat, sexe, educació musical, etc. Veieu amb més profunditat cadascun d'aquests apartats:

1^a part

Instruccions inicials:

En les instruccions inicials es presenten una sèrie de característiques i requisits necessaris per a un correcte funcionament i realització del qüestionari. Aquestes són les característiques més destacades:

- *Duració aproximada dels fragments de 30 segons:* es considera un temps suficient per a la consolidació en la percepció de les emocions del fragment, i no massa llarg com per què es perdi homogeneïtat en aquest.

- *S'ha de respondre segons les emocions que expressen els fragments, no les emocions que indueixen:* en la recerca es cerca una homogeneïtat entre les respostes del mateix fragment dels individus participants; preguntar per les emocions que expressen els fragments afavoreixen aquesta homogeneïtat, ja que redueix els efectes dels mecanismes típics en la inducció d'emocions (veieu la secció 2.5.).

- *S'avaluaran els graus d'alegria, tristesa, sorpresa, irritació, despreci i amor, així com si les emocions que es transmeten són positives, negatives, animades o relaxades.* La selecció d'emocions s'ha realitzat a partir de l'enfoc dimensional (veieu la secció 2.2.), i més concretament el model circumplex de Russell, que utilitza les dimensions d'activitat i valència per a la representació de les emocions dins un espai bidimensional. Així, a partir d'aquestes dues variables hem escollit emocions que puguin representar la majoria de posicions en l'espai d'aquest model. Les preguntes envers si les emocions són positives, negatives, animades o relaxades tenen relació directa amb els conceptes de valència i activitat respectivament. Veieu la figura 8:

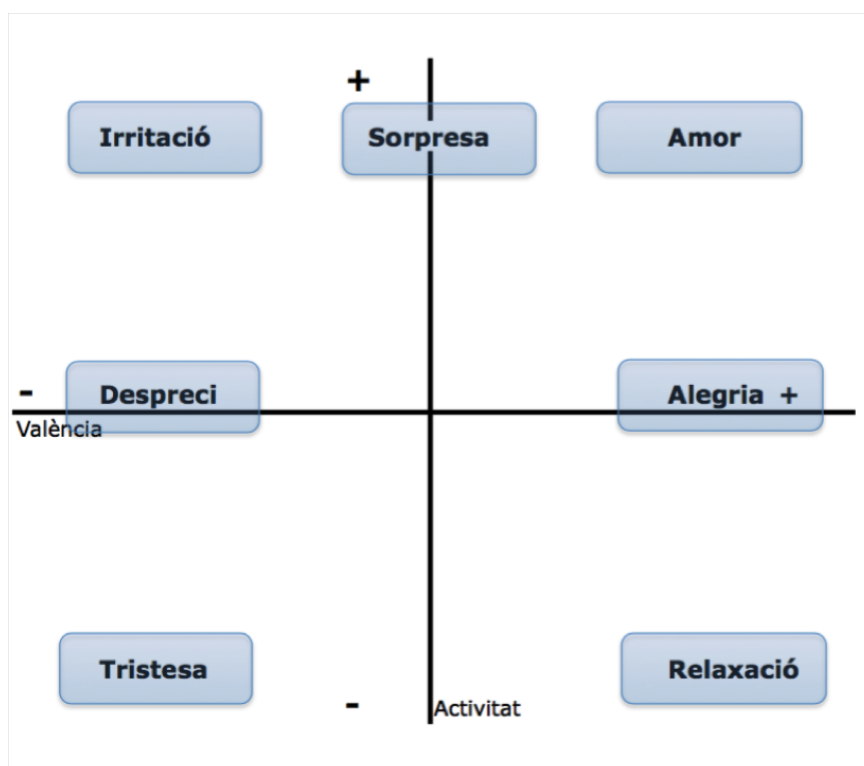


Fig. 8 Classificació d'emocions segons el mètode circumplex de Russell

- Respostes en una escala del 0 al 6: Es considera una gradació suficient que permet a l'individu tenir certa flexibilitat en les avaluacions, amb dues gradacions intermitges entre els extrems (0 i 6) i el punt intermig (3), però no massa gran perquè aquest perdi consciència de la intensitat de cadascun dels punts.

- El qüestionari consta de 30 fragments musicals. Tot i que la recerca conté 240 arxius o fragments diferents, cadascun dels individus n'avalua 30, limitant així la duració del qüestionari a 30 minuts i assegurant un estat de concentració suficientment alt com per què les respostes siguin fiables i amb la màxima atenció possible. Així, l'script que conté el qüestionari s'ha dissenyat per què reparteixi aquests 240 fragments entre els diferents usuaris de forma automàtica, mitjançant la programació en llenguatge *javascript* de 8 blocs diferents de 30 fragments cadascun; les respostes de cadascun dels blocs s'ordenen de forma aleatòria automàticament per a

controlar possibles efectes sistemàtics d'ordre i d'aprenentatge a mida que es va responent al qüestionari.

Model qüestionari

Les emocions en la música forma part del projecte de final de grau sobre la percepció de les emocions en la música dut a terme per Víctor Núñez Tarifa, estudiant de grau a l'Escola Superior de Música de Catalunya.

En el qüestionari a continuació et presentem una sèrie de fragments musicals d'uns 30 segons, a partir dels quals hauràs de respondre a una sèrie de preguntes.

Les preguntes a respondre tenen relació **amb les emocions que transmeten aquests fragments, i NO les emocions que sents o que t'indueixen a tu al escoltar-los**. Així doncs, a més d'avaluar els graus d'alegria, tristesa, sorpresa, irritació, despreci i amor que transmet cadascun dels fragments, hauràs de valorar en quin grau les emocions que transmeten són positives, negatives, animades o relaxades, responent a les preguntes en una **escala** que anirà del **0 (absència d'emoció) al 6**.

Cal tenir en compte que no hi ha cap resposta que sigui més bona que una altra, sempre i quant paris atenció a cadascun dels fragments i contestis a cadascuna de les respostes amb la major concentració possible. **En qualsevol moment pots parar, descansar i continuar el qüestionari més tard, sempre i quan no tanquis la finestra actual.**

El qüestionari consta de 30 fragments musicals, i té una duració aproximada de 30 minuts.

Pregunta realitzada per a cada arxiu (en aquesta pregunta, els participants mouen un slider per fer l'anotació)

_Si us plau, reproduiu l'àudio tantes vegades com sigui necessari i responeu a les preguntes a continuació

- 1.- Transmet **alegria**?
- 2.- Transmet **tristesa**?
- 3.- Transmet **sorpresa**?
- 4.- Transmet **irritació**?
- 5.- Transmet **despreci**?
- 6.- Transmet **amor**?
- 7.- Transmet **animació**?
- 8.- Transmet **relaxació**?
- 9.- Transmet **emocions positives**?
- 10.- Transmet **emocions negatives**?

2ª part:

Una vegada realitzades les preguntes a partir dels arxius, es realitzen preguntes complementàries que, com ja hem comentat anteriorment, poden servir per crear grups, establir relacions amb les respostes o extreure conclusions a partir de la comparativa de les respostes en els diferents grups creats en el cas que fos convenient.

Pregunta nº1 (resposta única):

De quin gènere ets?

- 1.- Masculí
- 2.- Femení

Pregunta nº2 (en aquesta pregunta, els participants mouen un slider per fer l'anotació):

- 1.- Quina edat tens?

2.- Durant quants anys has rebut formació musical?

3.- Quants en portes tocant un instrument?

Pregunta nº 3 (resposta única):

Quin instrument toques?

- 1.- Veu cantada
- 2.- Vent-metall / Saxo
- 3.- Vent-fusta (excepte Saxo)
- 4.- Violí / viola
- 5.- Violoncel
- 6.- Contrabaix
- 7.- Piano
- 8.- Baix elèctric
- 9.- Sintetitzadors / electrònica
- 10.- Percussió / bateria
- 11.- Guitarra clàssica / acústica
- 12.- Guitarra elèctrica
- 13.- Altres
- 14.- Cap

Pregunta nº 4 (resposta única):

Generalment escoltes música fixan-te en les emocions que hi transmet, o et fixes en altres aspectes?

- 1.- Principalment em fixo en els aspectes emocionals
- 2.- Em fixo en aspectes emocionals, però més en altres
- 3.- Em fixo tant en aspectes emocionals com en altres
- 4.- Em fixo poc en aspectes emocionals, i molt en altres
- 5.- No em fixo gens en els aspectes emocionals

Pregunta nº 5 (en aquesta pregunta, els participants mouen un slider per fer l'anotació):

En quin grau creus que la música t'emociona?

Pregunta nº 6 (en aquesta pregunta, els participants mouen un slider per fer l'anotació):

Indica quin grau de familiaritat teniu amb cadascun d'aquests estils:

- 1.- Jazz
- 2.- Pop
- 3.- Electrònica / Dance
- 4.- Metal
- 5.- Rock clàssic / Rock alternatiu
- 6.- Hip Hop
- 7.- Música predominantment vocal
- 8.- Música predominantment instrumental

4.3.- Anàlisi estadístic descriptiu de les dades obtingudes en el qüestionari

L'anàlisi de les dades és un pas imprescindible en la recerca. En aquesta part del procés s'estudien les característiques essencials de les dades i, mitjançant una sèrie de càlculs estadístics, es descriuen les tendències i comportaments observats, tant de forma numèrica com gràfica.

L'anàlisi es realitza amb els programes de càlcul *Microsoft Excel* i *R*, i consta de les següents fases:

- *Organització de les dades*. Per tal de facilitar la manipulació de les dades i la compatibilitat entre programes de càlcul, les anotacions dels subjectes es recullen en un sol arxiu de text en format csv (comma separated values).

- *Realització dels càlculs pertinents*. Una vegada manipulades i organitzades les dades de la forma pertinent, es realitzen els càlculs estadístics necessaris per a poder donar una descripció numèrica numèrica de les tendències i comportaments de cadascun dels aspectes a analitzar. Aquests càlculs són: moda, mediana, promig, desviació estàndard, desviació interquartil, anàlisi de la variança i mètode de Tukey. Vegem-los amb més detall:

- *Moda*: la moda és el valor amb major freqüència en un conjunt de dades
- *Mediana*: és el número central d'un grup de números ordenats per tamany.
- *Promig*: resultat de la suma de tots els valors d'un conjunt de dades, dividit pel nombre de conjunt de dades.
- *Desviació estàndard*: mesura de dispersió de les dades en valors continus
- *Desviació interquartil*: mesura de dispersió de les dades en valors discrets
- *Anàlisi de la variança (ANOVA)*: l'anàlisi de la variança és una tècnica estadística que permet mesurar l'efecte d'una sèrie de variables independents sobre una variable dependent; aquest efecte es mesura

al comparar les mitges de la variable dependent en diferent nivells dels factors. La hipòtesi nul.la estableix que totes les mitges de la població (mitges dels nivells dels factors) són iguals ($p > 0.05$), mentre que la hipòtesi alternativa estableix que almenys una és diferent ($p < 0.05$).

- *Mètode de Tukey*: també conegut com el mètode de comparacions múltiples, aquest mètode s'utilitza per determinar l'anàlisi de la variança de totes les parelles possibles de factors de la variable dependent.

- *Realització dels gràfics necessaris i anàlisi de les dades a partir dels resultats obtinguts.*

4.4. - Extracció de descriptors per cadascun dels arxius musicals de la col.lecció

Per aquesta tasca s'utilitzen dos tipus d'aplicatius: *PRAAT*⁴ (Boersma i Weenink, 2017) i *Essentia*⁵ (Bogdanov et al., 2013). PRAAT és un software per l'anàlisi científic de la parla, utilitzat també en recerca musical quan el que es vol analitzar són les característiques de la veu cantada; Essentia, en canvi, és una llibreria creada per l'anàlisi i extracció d'informació envers arxius d'àudio musicals.

L'extracció de descriptors és una tasca laboriosa, sobretot si estem parlant d'analitzar els descriptors musicals i vocals de 240 fragments musicals; per aquest motiu, és necessari que cadascun d'aquests processos sigui automatitzat amb *scripts* que facilitin aquesta tasca.

⁴ <http://www.fon.hum.uva.nl/praat/>

⁵ <http://essentia.upf.edu/documentation/>

Els descriptors escollits per la recerca han estat els següents:

Descriptors musicals (Essentia)

- *Danceability*: es tracta d'un algorisme que indica el grau en què la música és ballable. Com més alt el coeficient, més ballable és, i a l'inrevés.
- *Onset rate*: es pot considerar com la mesura del nombre d'events sonors per segon; un valor alt indica una densitat rítmica alta.
- *Dynamic Complexity*: és la desviació promitja absoluta respecte al nivell de sonoritat global; està relacionada amb el rang dinàmic o la quantitat de fluctuació del nivell de sonoritat en un enregistrament.
- *Beats sonoritat*: és una mesura de força de les pulsacions en una peça, i s'utilitza per caracteritzar una peça musical.
- *Barkbands*: és un algorisme que divideix l'espectre en 28 bandes Bark. Les divisions de freqüència són les següents (en Hz): 0, 50, 100, 150, 200, 300, 400, 510, 630, 770, 920, 1080, 1270, 1480, 1720, 2000, 2320, 2700, 3150, 3700, 4400, 5300, 6400, 7700, 9500, 12000, 15500, 20500 i 27000
- *Spectral centroid*: aquesta mesura indica el centre de gravetat de l'espectre, i té una connexió directa amb la brillantesa del so.
- *Spectral kurtosis*: la kurtosi és una mesura de la planura de la distribució envers el seu valor promig. Una kurtosi negativa indica un espectre pla, i una positiva un espectre amb més pics, i s'utilitza per la caracterització del timbre.
- *Spectral flux*: mesura com de ràpid l'energia de l'espectre varia, i s'utilitza per determinar el timbre d'una senyal d'àudio.
- *Spectral skewness*: és una mesura d'asimetria de la distribució respecte al seu valor promig. Un valor negatiu indica més energia en freqüències altes, i a l'inrevés.
- *Spectral rolloff*: és la freqüència en la qual per sota d'ella conté el 85% de l'energia de l'espectre. Aquesta mesura és utilitzada per distingir sons harmònics i sons amb soroll.
- *Energy Band Ratio (20-150 Hz), (150-800 Hz), (800-4000 Hz), (4000-20000 Hz)*: aquestes mesures indiquen la ràtio d'energia en cadascuna de les bandes indicades.
- *MFCC*: usats en el reconeixement de la veu, són una representació compacta de

l'envolopant espectral.

- *Pitch salience*: ve donada per la ràtio entre el pic més alt i el més baix en la funció d'autocorrelació. Els sons sense pitch tenen un valor de pitch salience proper a 0, mentre que els sons harmònics tenen valors propers a 1; aquesta mesura és utilitzada per caracteritzar sons percussius.
- *Silence rate -60 dB*: és el nombre de frames on el nivell està per sota de -60 dB, i és utilitzat per mesurar el nivell de compressió de la senyal.
- *Zero crossing rate*: és el nombre de canvis en la senyal entre valors de senyal consecutius i el valor total, i s'utilitza per mesurar el soroll d'una senyal.
- *Spectral complexity*: la complexitat espectral és una mesura de la complexitat en la instrumentació d'una peça musical.
- *Chord changes rate*: és la ràtio entre el nombre de canvis tonals d'acords i el nombre total d'acords detectats.
- *Chord number rate*: és la ràtio del nombre d'acords diferents tocats més d'un 1% del temps del nombre total d'acords detectats en la progressió tonal d'acords.
- *Chord strength*: representa la força/correlació en la detecció d'acords; un valor alt indica que la detecció de l'acord és molt tonal, i un valor baix indica una baixa detecció tonal.
- *Dissonance*: és un descriptor perceptual utilitzat en la mesura de la dissonància d'un so.
- *Key strength*: representa la força/correlació de la tonalitat; un valor alt indica que la peça és molt tonal, i a l'inrevés.
- *Thpcp*: és l'Harmonic Pitch Class Profile (HPCP) transposat. L'HPCP és un vector dimensional de 12 punts, cadascun dels quals representa un bin en una escala amb temperament igual.
- *Average sonoritat*: aquest algorisme computa el valor de sonoritat promig de la senyal.
- *Beats count*: computa el nombre de pulsacions de la senyal.
- *Bpm*: és la mesura de tempo del fragment.
- *Mean F0*: és el valor promig de la freqüència fonamental.
- *nPulses*: mesura el nombre de cops glotals.
- *pctUnvoiced*: mesura el nombre de frames de la senyal sense veu
- *pctVoicebreaks*: mesura la duració total de les parades entre parts amb veu en la senyal, dividides pel total de la duració d'aquesta.

- *Jitter (loc i rap)*: mesura el caràcter d'aperiodicitat de la vibració cordal; un valor alt s'associa a una percepció vocal disfònica, aspresa vocal i un elevat grau d'escapament aeri (breathiness).
- *Shimmer_loc*: mesura les variacions d'amplitud mitja entre cicles; és anàlog al Jitter però per la freqüència: està relacionat amb el grau de disfonia i amb la percepció de l'aire (breathiness).
- *HNR (Harmonic-to-noise ratio)*: aquest paràmetre estableix una relació entre l'energia acústica dels harmònics i el soroll generat per l'aire;
- *Autocorrelation*: Relacionat amb l'HNR, indica la correlació entre l'energia dels harmònics i el soroll; un valor baix indica una gran presència de soroll, i a l'inrevés.
- *H1 (primer formant)*. S'associa principalment a l'obertura mandibular i a l'altura de la llengua.
- *H2 (segon formant)*. s'associa a l'anteriorització i posteriorització de la llengua; com més anterior, el valor és més elevats.
- *H3 (tercer formant)* s'associa a llavis estirats o arrodonits; com més elevat el valor, llavis més estirats.
- *BDWH1, BDWH2, BDWH3*: s'associa amb la configuració del tracte vocal.
- *Hammerberg Index*: mesura la diferència d'energia espectral entre la banda de 0 a 2000 Hz i de 2000 Hz a 5000 Hz.

4.5.- Creació de models

En aquesta part del procés es cerquen correlacions entre la variable dependent (les emocions i les anotacions dels subjectes) i les variables explicatives (els descriptors) mitjançant una sèrie de tècniques d'aprenentatge automàtic, escollides en funció de les característiques de les dades a analitzar i de l'objecte a analitzar.

Una vegada realitzada aquesta correlació es realitzen models de predicció mitjançant la tècnica de cross-validació de 10 parts, que consisteix en dividir el conjunt de dades del que disposem en deu parts, de les quals nou serveixen per entrenar al model i una per mesurar la precisió de la predicció.

El procediment s'ha realitzat amb Weka⁶ (Witten, I.H. i Frank, E.,2005), un software de mineria de dades amb una col.lecció d'algorismes per tasques d'aprenentatge automàtic desenvolupat per la Universitat de Waikato. Aquest consta de les següents fases:

- *Manipulació de les dades.* Per tal de poder ser utilitzades dins l'entorn de Weka, les dades han de presentar-se en un únic arxiu de text en format csv, amb les columnes de les variables explicatives a l'esquerra i les de les variables dependents a la dreta, ja que el programa interpreta de forma automàtica que les columna més a la dreta correspon a la variable dependent (es pot seleccionar una columna diferent una vegada carregat l'arxiu).

- *Definició de descriptors rellevants.* En la definició de descriptors s'estableixen les correlacions entre els descriptors i les anotacions dels subjectes; pel càlcul de les correlacions s'utilitzarà la tècnica d'aprenentatge automàtic *support vector machines*.

- *Realització de models de predicció.* A partir de les correlacions establertes en el procediment de definició de descriptors, es realitza la cross-validació de 10 parts; a partir dels resultats d'aquest procediment es podrà avaluar l'efectivitat de cadascun dels models.

4.6.- Reproductibilitat

La reproductibilitat es refereix a la capacitat d'una prova o experiment de ser reproduït o replicat per altres, en particular per la comunitat científica.

Per això és necessari que el material creat en cadascun dels processos de la recerca estigui disponible per la comunitat, amb el material i la documentació detallada de les instruccions necessàries per a la seva reproducció.

Així doncs, el material és inclòs a la plataforma de desenvolupament de software

⁶ <http://www.cs.waikato.ac.nz/ml/weka/>

GitHub⁷, amb els següents arxius adjunts:

- *Emo_database.zip*: arxiu comprimit de la col·lecció d'àudios utilitzats en la recerca, amb un arxiu per cada fragment i condició.
- *PsyToolkit_script.txt*: el model de l'script utilitzat per a la realització del qüestionari, en format .txt
- *CSV1.csv i 5CSV1.csv*: arxius en .csv (comma separated values) de les dades obtingudes a partir de les respostes dels subjectes al qüestionari.
- *Essentia_com.txt*: model de les instruccions donades en línia de comandes a la llibreria Essentia per a l'extracció dels descriptors de la música.
- *Praat_script.txt*: model de l'script utilitzat per l'automatització del procés d'extracció de descriptors de la veu
- *Weka_Dades.csv*: arxiu en .csv que reuneix les anotacions dels subjectes i els descriptors utilitzats en la recerca.
- *Readme.txt*: full informatiu de les instruccions a seguir per a la reproductibilitat de l'experiment

⁷ <https://github.com/>

5.- Resultats

En aquesta secció es presenten els resultats de cadascun de les anàlisis realitzats en el procés de modelatge d'emocions; així doncs, primerament s'ha realitzat un anàlisi estadístic descriptiu de les dades obtingudes a partir de les anotacions dels subjectes, i a continuació s'han estudiat els resultats de la creació de models.

5.1.- Anàlisi estadístic descriptiu de les anotacions

L'anàlisi estadístic descriptiu consta de les següents parts: a) anàlisi dels subjectes, b) anàlisi de les emocions sense diferenciar condicions i c) anàlisi de les emocions diferenciant condicions.

Per poder realitzar l'anàlisi de les dades ha estat necessari recodificar les anotacions dels subjectes de 7 punts a 5 punts, ja que una excessiva dispersió de les dades dificultava la detecció dels models de comportament i característiques més rellevants (veieu Annex Hist_Subj_7 de l'Annex A); la recodificació s'ha realitzat agrupant els valors de l'escala de 7 punts d'aquesta forma: 0 (1 2) 3 (4 5) 6.

5.1.1. Anàlisi dels subjectes

L'anàlisi dels subjectes és un element molt important en la descripció de resultats de la recerca, ja que ens permet observar les tendències d'aquests en les seves anotacions i ens permet identificar possibles irregularitats, com subjectes amb un alt percentatge d'anotacions amb el mateix valor, o fragments o emocions en les quals els subjectes mostrin ambigüitat en les seves respostes.

El primer que intentarem detectar a partir de les anotacions de les que disposem és si els subjectes participants han realitzat les seves respectives proves amb atenció o, en canvi, han realitzat la prova de forma aleatòria o sense prestar atenció. Per analitzar aquestes possibles irregularitats es proposen dos tipus de proves: la primera analitza si

els subjectes han respòs a les preguntes utilitzant de forma excessiva el mateix valor, amb el que podria indicar un intent per respondre de forma automàtica i sense atenció a les preguntes del qüestionari; la segona intenta detectar canvis radicals en els valors de les anotacions de la primera part del qüestionari envers l'última part, suposant així un canvi en l'interès del subjecte per realitzar la prova.

A la primera prova es realitzen histogrames de les anotacions de cada subjecte i es seleccionen els subjectes en els quals un valor predomina de forma excessiva. Si observem els histogrames Hist_Subj_5 de l'Annex A, podem veure com hi ha una sèrie de subjectes que han utilitzat el valor 0 en una proporció molt gran. Aquests són els subjectes 10, 16, 33 i 38. Observem ara les taules de freqüència per cadascun d'ells:

Subjectes	0	1	2	3	4
10	85,0	NA	6,7	6,0	2,3
16	70,3	21,0	1,7	5,7	1,3
33	71,3	16,3	8,7	3,7	NA
38	74,3	17,3	3,3	2,3	2,7

Taula 9. Percentatge dels valors respecte al total de cadascun dels subjectes indicats a la primera columna

La taula 9 mostra un excés de respostes amb el valor 0 en cadascun dels subjectes indicats, principalment en els casos dels subjectes 10 i 38; el que fa suposar un desinterès d'aquests per la realització de la prova; d'aquesta forma, i per garantir la fiabilitat de les dades, els subjectes 10 i 38 són descartats per als posteriors anàlisi i la creació de models.

Per a la realització de la segona prova de detecció d'irregularitats, es comparen els histogrames corresponents al primer terç i al tercer terç del global de respostes de cadascun dels subjectes i s'intenten detectar possibles canvis radicals entre ells. Si observem els histogrames Subj_Glob_Ini i Subj_Glob_Fin de l'Annex B, podem veure que no hi han canvis de tendència dràstics en cap dels subjectes participants.

Fins ara ens hem trobat amb una sèrie de subjectes que tendeixen a avaluar els respostes amb un percentatge molt alt en el valor 0; veiem si aquesta tendència es manté en la majoria de subjectes del qüestionari:

0	1	2	3	4
35,3	26,7	11,5	22,5	4

Taula 10. Percentatge dels valors de tots els subjectes respecte al total

Com es mostra en la taula 10 les anotacions dels subjectes tendeixen a concentrar-se en els valors 0, 1 i 3. Veiem si aquesta tendència es compleix en la majoria de subjectes:.

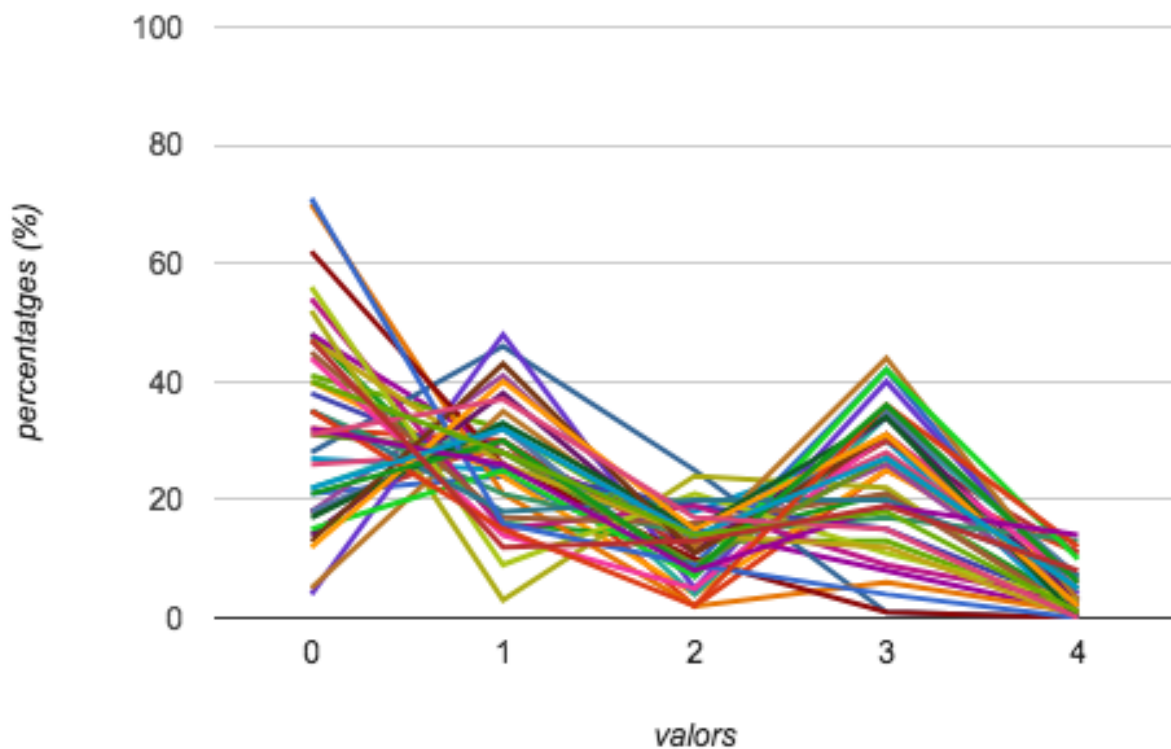


Fig 11. Percentatge dels valors, per cadascun dels subjectes

A la Fig 11 es pot observar com els subjectes tendeixen a avaluar les seves respostes amb valors de 0,1, i 3; no obstant, el valor 0 presenta força dispersió de les dades.

L'última prova en l'anàlisi de subjectes pretén estudiar possibles ambigüitats en les anotacions de les emocions per cadascun dels fragments, l'ambigüitats en les dades obtingudes. Aquestes ambigüitats es donen en emocions sense una tendència clara (valors centrals propers a 2) amb valors repartits (mateixa proporció de 0,1,2,3 i 4) o bipolaritzats (predominança de valors en els extrems). Per aquesta prova es realitzen histogrames de cadascun dels ítems i es seleccionen els histogrames en els que els valors tinguin una distribució similar o bipolar en l'eix horitzontal (Veieu els histogrames Hist.items_x de l'Annex C); com a suport als histogrames es realitzen els càlculs de la mediana i el primer i tercer quartil, que reafirmen la posició d'ambigüitat si el valor de la mediana és proper a 2 i el primer i tercer quartil estan allunyats d'aquesta a una distància similar (Veieu Càlculs de l'Annex C).

Els resultats de les anàlisis mostren una sèrie d'ítems ambigus que indiquen una certa dificultat dels subjectes a l'hora de percebre aquella emoció en aquell fragment particular. La taula 12 recull els percentatges dels ítems aparentment ambigus per cadascuna de les emocions:

Alegria	Tristesa	Sorpresa	Irritacio	Despreci
12,8	7,7	15,4	7,7	7,7
Amor	Emocposit	Emocnegat	Animacio	Relaxacio
7,7	12,8	5,1	10,3	12,8

Taula 12. Percentatge d'ítems ambigus respecte al total d'aquests

La taula 12 mostra un percentatge alt d'ambigüitat en l'emoció sorpresa, mentre que en l'emoció negativa, en canvi, el percentatge és força baix; a la vegada, si s'analitzen els percentatges de forma més àmplia, podem observar com les emocions amb una valència positiva presenten un grau d'ambigüitat més alt que les emocions amb valència

negativa (excepte Amor).

Si observem de nou els ítems seleccionats (veieu Càlculs de l'Annex C) podem veure com els fragments 33 i 52 tenen tres ítems seleccionats cadascun; això indica amb força claredat que en aquests fragments les emocions es perceben amb més dificultat; donat l'alt grau d'ambigüïtat d'aquests fragments, s'ha decidit que siguin descartats en els anàlisi posteriors i en la creació de models.

5.1.2. Anàlisi de les emocions sense diferenciar condicions

En aquesta part de la definició dels resultats es descriuen els comportaments de les emocions sense tenir en compte les condicions; per tal de realitzar aquest estudi, s'han reduït les dades de forma que els judicis de cadascuna de les condicions s'han unit per establir una condició global.

L'anàlisi de les emocions sense diferenciar condicions ens permet observar de forma global les tendències de cadascuna de les emocions al llarg de la col·lecció d'àudio; per aquest anàlisi realitzarem dos tipus de proves: la primera ens indicarà si les emocions compleixen els mínims necessaris per a ser considerada útils i representatives en la recerca, i la segona ens permetrà extreure informació envers quines emocions es perceben amb més freqüència a partir dels percentatges de les puntuacions donades a cadascuna de les emocions.

Tenint en compte els criteris i procediments utilitzats en l'àmbit de la ciència afectiva, per a què una emoció pugui ser considerada vàlida i representativa és necessari que hagi estat avaluada amb valors diferents de 0 com a mínim en una quarta part (25%) del total de la mostra de la que es compon la recerca; la taula 4 mostra el percentatge de fragments amb valors diferents de 0:

Alegria	Tristesa	Sorpresa	Irritacio	Despreci
89,7	87,9	77,6	46,6	31,0
Amor	Emocposit	Emocnegat	Animacio	Relaxacio
79,3	91,4	74,1	91,4	81,0

Taula 13. Percentatge dels valors diferents de 0 per cada emoció

La taula 13 mostra com totes les emocions compleixen el mínim per ser considerades vàlides.

En la segona prova, com ja hem dit, es descriu el comportament de cadascuna de les emocions a partir dels resultats obtinguts de les anotacions dels subjectes per cada emoció, podent observar així quines emocions es perceben amb més o menys freqüència; per l'anàlisi s'ha realitzat un histograma i un gràfic amb els percentatges dels resultats obtinguts per cadascuna de les emocions. Veiem doncs la taula 14 i la Fig 15:

Alegria	Tristesa	Sorpresa	Irritacio	Despreci
10,8	12	8	8	6,3
Amor	Emocposit	Emocnegat	Animacio	Relaxacio
10,1	12,6	9,3	14,7	8,2

Taula 14. Percentatge dels resultats de les anotacions, per cadascuna de les emocions

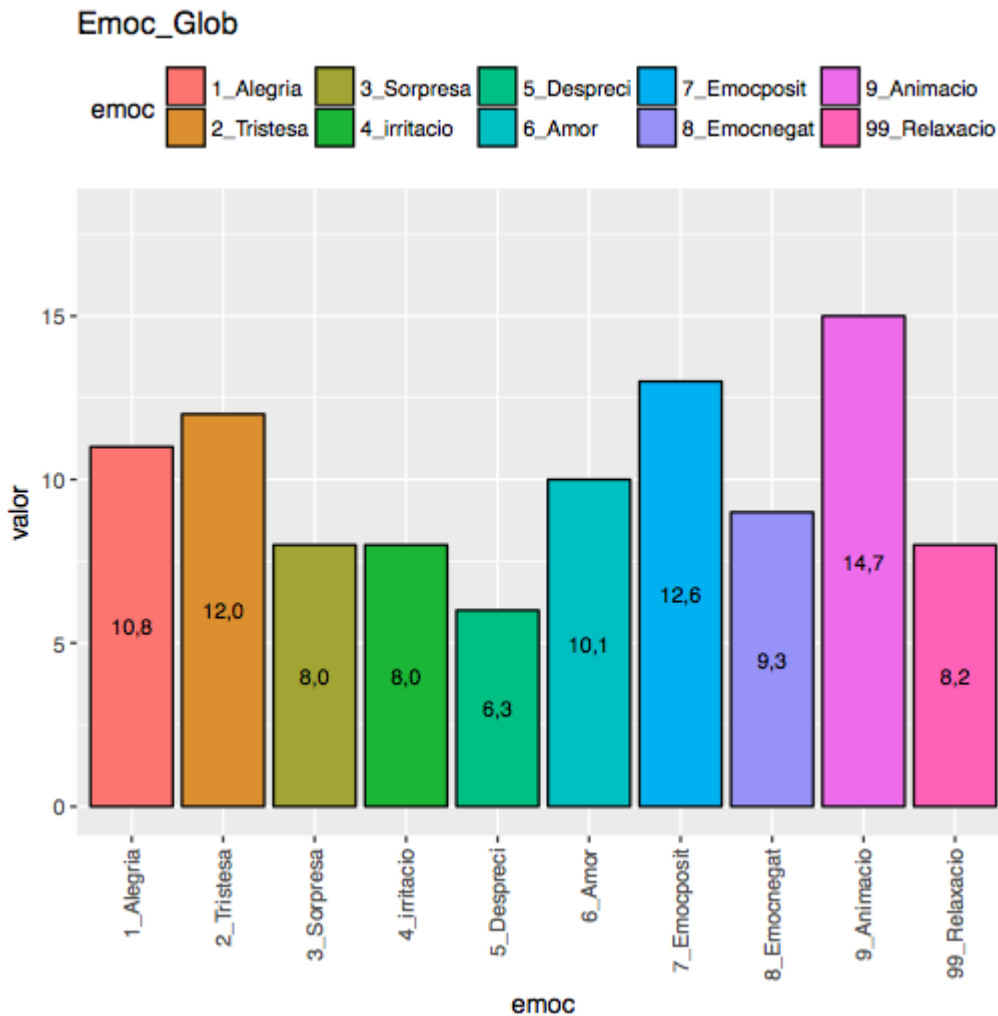


Fig 15. Percentatge dels resultats de les anotacions, per cadascuna de les emocions

La taula 14 i l'histograma de la Fig 15 mostren com l'emoció "Animació", és la que es percep amb més freqüència i "Despreci", en canvi, és la que menys; d'una altra banda, si observem els resultats d'una forma més àmplia, podem veure com, per norma general, les emocions amb valència positiva tendeixen a ser més detectables que les que tenen valència negativa (excepte "Tristesesa"); a la vegada, no hi ha una tendència clara envers la relació de les valoracions amb el seu grau d'activitat.

5.1.3. Anàlisi de les emocions diferenciant condicions

En aquesta última part de l'estadístic descriptiu s'estudien els comportaments de les emocions tenint en compte les diverses condicions proposades en la recerca: Veu, Instrumental, Veu + Instrumental i Mescla.

El primer que intentarem detectar en aquesta part de l'anàlisi és l'efecte de les condicions sobre les anotacions dels subjectes sense diferenciar emocions, de forma que ens permeti donar una idea de quines condicions perceben amb més freqüència les emocions. Veieu la taula 16:

Veu	Inst	Veulnst	Mix
24,4	23,1	26	26,5

Taula 16. Percentatge dels judicis diferents de 0, per cadascuna de les condicions

En la taula 16 es mostra el percentatge de valors en cadascuna de les condicions respecte al total, dels quals es pot deduir que la condició que percep les emocions amb més freqüència és la de mescla, mentre que la condició instrumental les percep en menor mesura; observem si aquestes diferències són suficientment rellevants.

Per analitzar si les diferències entre els factors de la condició són suficientment rellevants, és necessari realitzar l'anàlisi de la varianza (ANOVA) d'aquesta variable; primer de tot, observem si la variable condició és un efecte rellevant en la percepció de les emocions observant els resultats del valor de significació en l'anàlisi de la varianza, i si ho és (el valor ha de ser menor a 0.05), realitzem una comparativa per parelles de cadascun dels seus factors amb el mètode de Tukey, el qual ens indicarà quines parelles són significatives.

Veieu la taula 17:

	Sum squares	Df	Mean square	F	Sig,
Cond	42	3	13,88	8592	0,000 ***
Residuals	17559	10869	1,62		

Taula 17 Resultats de l'anàlisi de la varianza ANOVA

L'anàlisi de la varianza de la taula 11 indica un efecte significatiu de la condició en la percepció d'emocions, $F(3,10869)=8.592, p=0.000$. Veiem ara si aquest efecte significatiu es mostra en la comparativa de totes les condicions o solament en alguna d'elles; per realitzar això s'utilitza el mètode de Tukey, el qual permet fer l'anàlisi de la varianza en totes les combinacions possibles de parelles de la condició:

	Difference	Lower	Upper	p-Value
Veü-Inst	0,073	-0,016	0,162	0,152
Veü-VeüInst	0,063	-0,025	0,152	0,257
Veü-Mix	-0,087	-0,175	0,002	0,057
Inst-VeüInst	0,136	0,047	0,225	0,000 ***
Inst-Mix	0,16	0,071	0,248	0,000 ***
VeüInst-Mix	-0,023	-0,112	0,065	0,903

Taula 18. Comparativa de les condicions sense diferenciar emocions a partir de l'anàlisi de la varianza

La taula 18 mostra mitjançant els asteriscs les parelles de condicions que per les seves diferències són estadísticament significatives ; així doncs, podem observar com les parelles Instrumental – Veü Instrumental i Instrumental – Mix mostren una diferència significativa en la percepció d'emocions, principalment deguda a la condició Instrumental (ja que és la comuna en les dues parelles afectades).

Fins ara s'ha realitzat un estudi de l'efecte de la condició sobre les emocions de forma global; anem a observar que succeeix quan diferenciem emocions.

Veiem els histogrames de la Fig. 19:

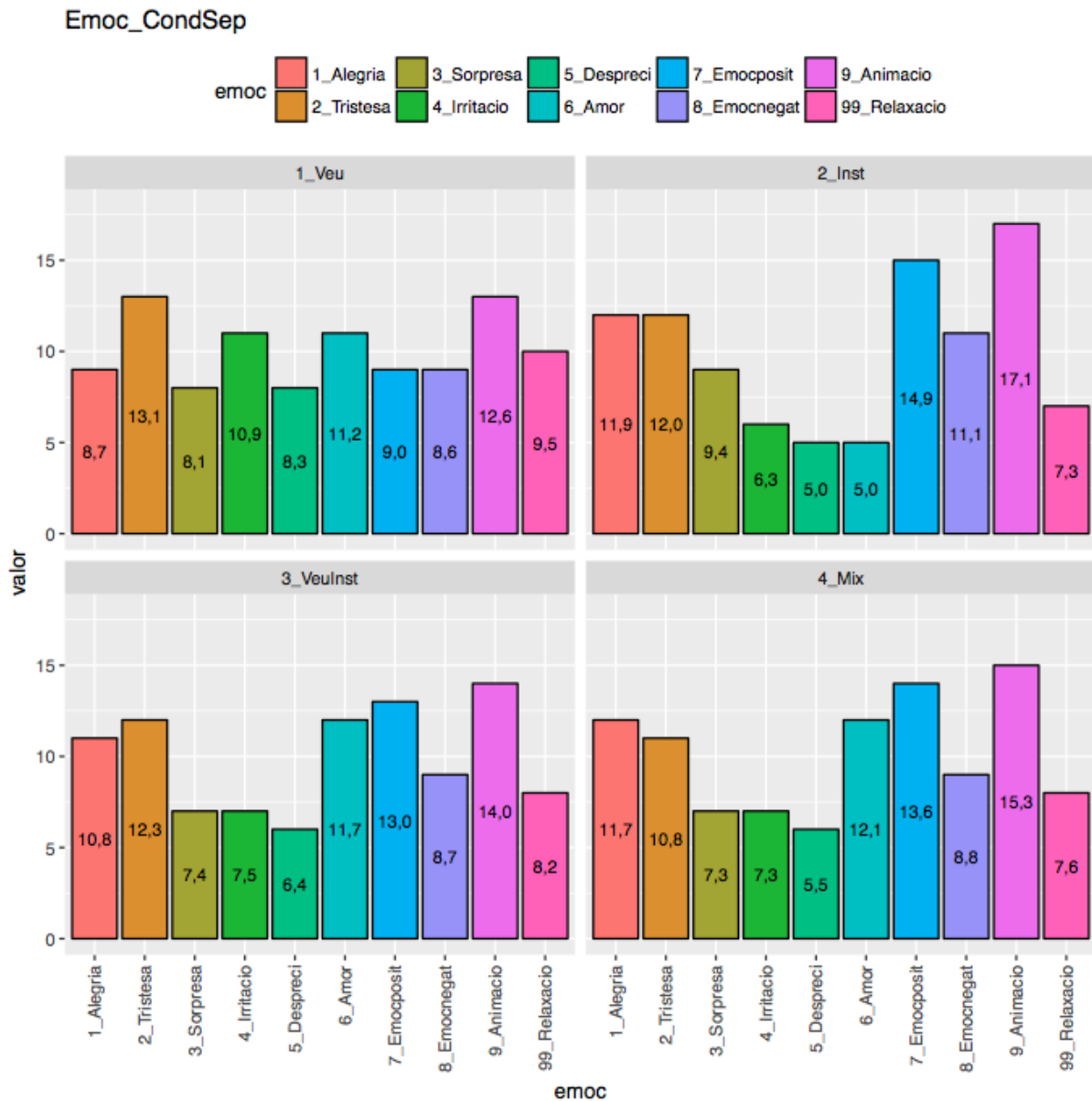


Fig 19. Percentatge del resultat de les anotacions, per cadascuna de les emocions i condicions

En els histogrames de la Fig 19 es mostra el percentatge de les valoracions en cadascuna de les condicions, les quals tenen les següents característiques:

- *Veu*: les emocions Tristesa i Animació es perceben amb més freqüència , mentre que Sorpresa i Despreci tenen els valors més baixos de percepció.
- *Instrumental*: les emocions Animació i Emoció Positiva tenen valoracions més altes, i les de Despreci i Amor mostren valors molt baixos.
- *Veu i Instrumental*: Animació i Emoció Positiva tenen valors més alts de percepció, mentre que Despreci i Sorpresa els tenen més baixos
- *Mescla*: Animació i Emoció Positiva es perceben amb més freqüència, i Despreci, Sorpresa i Irritació es perceben menys.

Un altre aspecte important en l'anàlisi de les emocions per condicions és la comparativa de les emocions i les seves tendències en funció de la condició utilitzada; per mostrar aquestes tendències s'han realitzat taules que mostren els percentatges de les emocions per condició respecte al total de valoracions de l'emoció, i gràfics que reafirmen aquestes dades. Veiem la taula 20:

	Veu	Inst	Veulnst	Mix
Alegria	19,7	25,4	26,1	28,8
Tristesa	26,6	23	26,7	23,8
Sorpresa	24,8	27	24	24,1
Irritacio	33,2	18,3	24,4	24,1
Despreci	32	18,3	26,4	23,3
Amor	27,1	11,4	29,9	31,6
Emocposit	17,5	27,2	26,8	28,6
Emocnegat	22,7	27,7	24,4	25,3
Animacio	20,9	26,8	24,6	27,6
Relaxacio	28,5	20,6	26	24,8

Taula 20. Percentatge dels resultats de les anotacions, per cadascuna de les emocions i condicions

Emoc_Cond

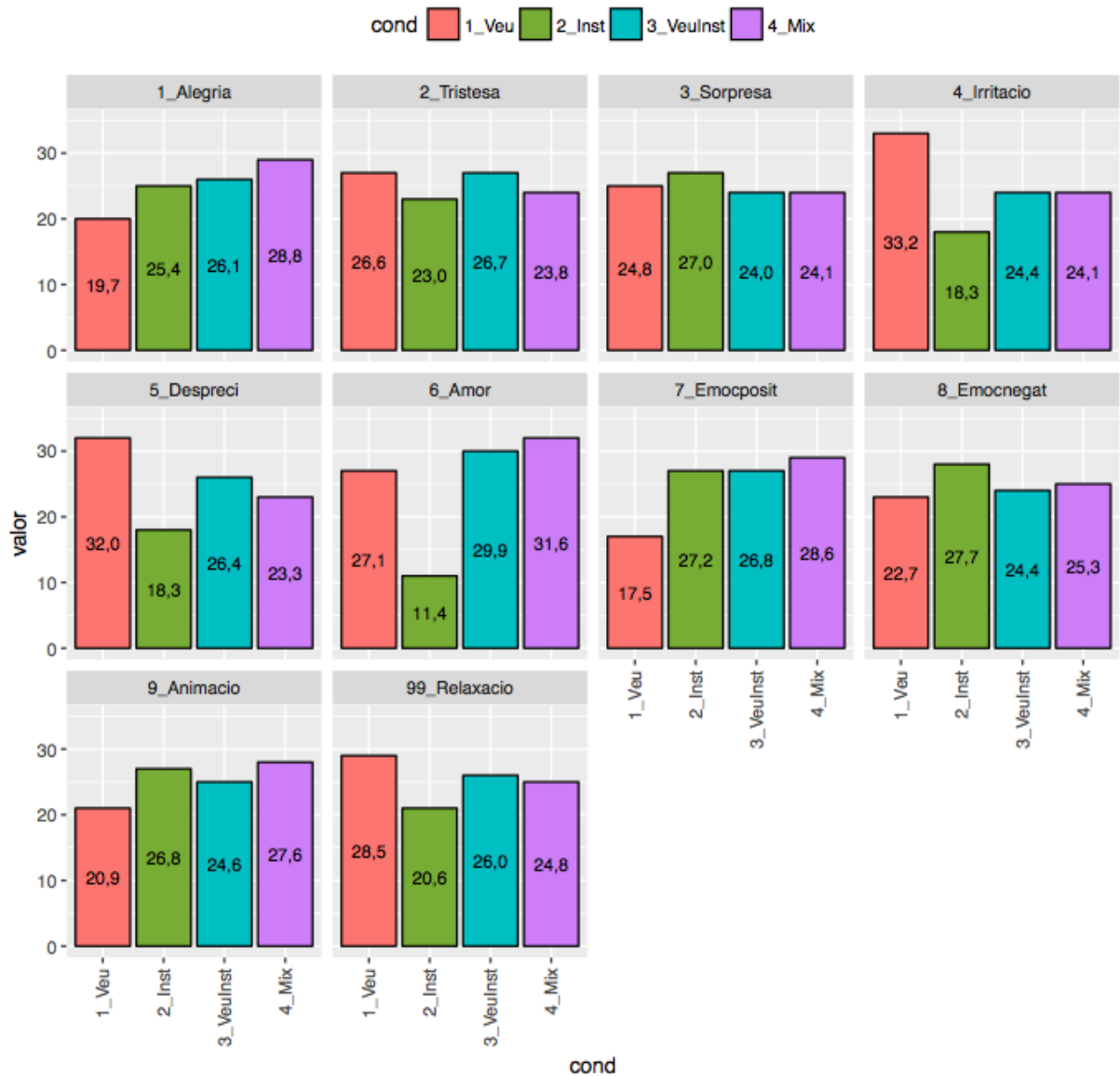


Fig 21. Percentatges de les puntuacions de les anotacions, per cadascuna de les emocions i condicions

La taula 20 i els histogrames de la Fig 21 mostren el percentatge de cadascun dels valors en cadascuna de les emocions i condicions, amb tendències força clares en cadascuna d'elles:

- *Alegria*: Té una tendència creixent, amb els valors més baixos en la condició de la Veu.

- *Tristesa*: Té una tendència decreixent excepte amb la condició Instrumental que presenta els valors més baixos.
- *Sorpresa*: Tendeix a ser estable en cadascuna de les condicions, malgrat que la condició instrumental presenta valors més alts.
- *Irritació*: Presenta un valor molt alt en la condició de la veu i molt baix en la condició instrumental.
- *Despreci*: Presenta un valor molt alt en la condició de la veu i molt baix en la condició instrumental.
- *Amor*: Presenta un valor molt alt en la condició de la veu i molt baix en la condició instrumental.
- *Emoció positiva*: Té una tendència creixent, amb els valors més baixos en la condició de la veu.
- *Emoció negativa*: Tendeix a ser estable en cadascuna de les condicions, malgrat que la condició instrumental presenta valor més alts.
- *Animació*: Presenta els valors més baixos en la condició de la veu i tendeix a ser estable en la resta.
- *Relaxació*: Presenta els valors més alts en la condició de la veu i tendeix a ser estable en la resta.

Així doncs, de forma global podem dir que les emocions “Alegria”, “Emoció positiva” i “Animació” (totes amb valència positiva i activitat no negativa) tendeixen a percebre’s amb més freqüència en la condició instrumental que a la de la veu, mentre que en les emocions “Irritació”, “Despreci”, “Amor” i “Relaxació”, tendeix a succeir a l’inrevés.

Veiem ara si l’efecte condició afecta sobre cadascuna de les emocions, o si solament és significatiu en alguna d’elles;, primer de tot realitzem l’anàlisi de la variança, i a continuació utilitzem el mètode de Tukey per estudiar amb més detall quines són les condicions estadísticament significatives. Veiem les taules 22 i 23:

	Sum squares	Df	Mean square	F	Sig
Alegria	46,40	3	15,46	10,92	0,000 ***
Tristesa	10,70	3	3,58	2,38	0,068
Sorpresa	2,40	3	0,79	0,59	0,623
Irritació	54,20	3	18,05	10,97	0,000 ***
Despreci	29,60	3	9,87	7,40	0,000 ***
Amor	204,20	3	68,07	42,95	0,000 ***
Emocposit	86,60	3	28,88	18,49	0,000 ***
Emocnegat	11,00	3	3,66	2,51	0,057
Animació	50,60	3	16,86	11,12	0,000 ***
Relaxació	15,80	3	5,28	4,06	0,007 **

Taula 22. Resultats de l'anàlisi de la variança per a cadascuna de les emocions (un valor de significació $p < 0.05$ indica una diferència entre mitges significativa).

Com mostra la taula 22, l'efecte condició afecta especialment a les emocions "Alegria", "Irritació", "Despreci", "Amor", "Emoció positiva" "Animació" i "Relaxació" mentre que les emocions de "Tristesa", "Sorpresa" i "Emoció negativa" no es veuen afectades. Veiem amb més detall quines són les parelles de condicions estadísticament significatives:

Índex (p)	Ve-u-Inst	Ve-u-Inst	Ve-u-Mix	Inst-Ve-u-Inst	Inst-Mix	Ve-u-Inst-Mix
Alegria	0.001**	0.000**	0.000**	0.999	0.311	0.369
Tristesa	0.209	0.994	0.428	0.125	0.973	0.289
Sorpresa	0.999	0.671	0.942	0.669	0.944	0.943
Irritacio	0.000**	0.002**	0.000**	0.164	0.281	0.990
Despreci	0.000**	0.166	0.011**	0.053*	0.420	0.736
Amor	0.000**	0.772	0.276	0.000**	0.000**	0.834
Emocposit	0.000**	0.000**	0.000**	0.929	0.957	0.676
Emocnegat	0.038*	0.810	0.823	0.282	0.258	0.999
Animacio	0.000**	0.005**	0.000**	0.520	0.954	0.234
Relaxacio	0.004**	0.321	0.078*	0.310	0.718	0.899

Taula 23. Comparativa de les condicions sense diferenciar emocions a partir de l'anàlisi de la variança ANOVA

A partir de la taula 23 podem observar com les parelles en les que apareixen diferències significatives varien en funció de l'emoció; així doncs podem dir d'una forma àmplia que en les emocions "Alegria", "Irritació", "Emoció positiva" i "Animació", les diferències significatives es troben en les parelles amb la condició de la Veu, mentre que en les emocions de "Despreci" i "Amor", es troben en les parelles amb la condició instrumental; en la resta no s'aprecien tendències significatives.

5.2.- Anàlisi de resultats de la creació de models

En aquest capítol es descriuen els comportaments i tendències observats a partir dels resultats obtinguts en la creació dels models, de forma que es pugui respondre a aspectes com quines emocions obtenen millors i pitjors ratings de predicció, quin model mostra uns valors més alts de predicció o quina influència tenen els components d'una cançó en els resultats.

El modelatge d'emocions s'ha realitzat a partir de dos tècniques diferents: la *regressió i la classificació*; la primera realitza una predicció a partir de valors continus de la variable dependent (l'emoció), i la segona discretitza els valors continus i els divideix en tres categories (absència d'emoció, resultat ambigu, presència d'emoció).

L'anàlisi consta de les següents parts: a) *anàlisi comparatiu del grau d'efectivitat dels models* b) *anàlisi de les emocions sense diferenciar condicions* c) *anàlisi de les emocions diferenciant condicions* d) *anàlisi dels resultats de la combinació de condicions*

5.2.1.- Anàlisi comparatiu del grau d'efectivitat dels models

En l'anàlisi comparatiu dels models es realitza una comparativa de l'efectivitat del model regressiu i el de classificació.

Com ja hem comentat anteriorment, el model de regressió cerca predir els valors de cadascuna de les emocions amb un valor numèric continu; el model de classificació, en

canvi, discretitza les dades de les anotacions obtingudes amb anterioritat i les categoritza en tres nivells diferents, de forma que la predicció es basa en resultats amb tres possibles solucions.

Un aspecte important en el model de classificació és la selecció dels intervals de valors en la categorització de les dades, ja que això definirà l'efectivitat o la inefectivitat d'aquest model.

Per a que un model sigui efectiu, és convenient que els valors de les anotacions estiguin repartits de forma uniforme entre categories, de forma que l'algorisme d'aprenentatge tracti a cadascuna de les categories amb el mateix grau d'importància.

Així doncs, la primera part de l'anàlisi comparatiu es centra en seleccionar l'interval de valors més adequat per la recerca; per a realitzar aquesta selecció es realitza una comparativa de dos models intervàlics diferents; el primer manté una distància entre intervals equidistants ($<0.33, 0.33-0.66, >0.66$) i el segon concentra més els valors baixos i dilata els alts ($<0.25, 0.25-0.55, >0.55$), de forma que aquests s'adequin més als resultats de l'anàlisi estadístic que mostra una tendència en l'avaluació de les respostes amb valors entre 0 i 1. Veiem la taula 24:

%	Interval 1	Interval 2
Alegria	53,4	47,8
Tristesa	53,4	48,3
Sorpresa	49,6	47,4
Irritacio	70,3	63,4
Despreci	74,6	66,8
Amor	64,7	53
Emocposit	56	42,7
Emocnegat	68,1	47,8
Animacio	56,9	50,9
Relaxacio	51,3	47,4

Taula 24. Comparativa dels models de classificació en discretitzacions amb intervals diferents

La taula 24 mostra amb claredat com el model amb les discretitzacions equidistants (Interval 1) ha obtingut uns valors d'efectivitat de la predicció majors que el model amb les discretitzacions amb un pes més gran en els valors baixos; degut a aquests resultats, el model interval 1 serà l'utilitzat en les posteriors anàlisi com a referència dels models de classificació.

Una vegada realitzada aquesta comparativa, estem en disposició d'observar què succeeix quan comparem els models regressius amb els de classificació. Veiem la taula 25:

%	Interval 1	Interval 2
Alegria	53,4	47,8
Tristesa	53,4	48,3
Sorpresa	49,6	47,4
Irritacio	70,3	63,4
Despreci	74,6	66,8
Amor	64,7	53
Emocposit	56	42,7
Emocnegat	68,1	47,8
Animacio	56,9	50,9
Relaxacio	51,3	47,4

Taula 25. Percentatges dels resultats de predicció per emocions en models de regressió i classificació i la diferència entre ells.

La taula 25 mostra els percentatges d'efectivitat de la predicció dels models de regressió i classificació en el global de les dades, així com la diferència entre ells; a partir d'aquestes dades podem observar amb claredat com els models basats en la categorització obtenen millors valors d'efectivitat en cadascuna de les emocions, mostrant diferències especialment gran en les emocions "Sorpresa", "Amor", "Animació" i "Relaxació", totes amb valència positiva.

5.2.2. Anàlisi de les emocions sense diferenciar condicions

L'anàlisi de les emocions sense diferenciar condicions permet estudiar el comportament i les tendències de les emocions en els models de predicció. Veieu la taula 28:

%	Regressió	Classificació
Alegria	43,23	53,4
Tristesa	36,02	53,4
Sorpresa	17,16	49,6
Irritacio	56,95	70,3
Despreci	46,06	74,6
Amor	33,25	64,7
Emocposit	41,88	56
Emocnegat	48,24	68,1
Animacio	21,41	56,9
Relaxacio	17,28	51,3

Taula 26. Percentatges dels resultats de predicció per emocions en models de regressió i classificació

La taula 26 mostra els percentatges dels resultats de la predicció en el model de regressió i el de classificació; donat que en la comparativa de models hem observat com en general el model de classificació dóna millors resultats d'efectivitat, analitzarem cadascun d'ells de forma independent, sabent que si no es diu el contrari, el model de classificació obté millors resultats:

- *Regressió*: En el model regressió, les emocions amb millors resultats en la predicció són l'emoció "Sorpresa" i "Relaxació", mentre que l'"Emoció negativa" i "Irritació" (ambdes amb valència negativa) obtenen els pitjors resultats.
- *Classificació*: En el model classificació, les emocions amb millors resultats en la predicció són l'emoció "Sorpresa" i "Relaxació", mentre que "Irritació" i "Despreci" (ambdues amb valència negativa) obtenen els pitjors resultats.

5.2.3.- Anàlisi de les emocions diferenciant condicions

En l'anàlisi de les emocions diferenciant condicions s'observen els resultats de l'efecte condició per determinar si aquesta és rellevant en la predicció d'emocions.

La primera prova de l'anàlisi consisteix en utilitzar la condició com a descriptor de les emocions i observar el comportament i tendència de les prediccions en cadascuna de les emocions; per poder realitzar la prova, s'inclou la condició com a descriptor, en el qual cadascuna de les condicions representa un número; així la condició veu representa el número 1, la condició instrumental el 2 i així successivament, i s'analitzen els resultats. Veieu la taula 29:

%	Regressió		Classificació	
	SI	NO	SI	NO
Alegria	43,76	43,23	52,6	53,4
Tristesa	37,01	36,02	53,4	53,4
Sorpresa	18,74	17,16	48,3	49,6
Irritació	57,55	56,95	68,1	70,3
Despreci	45,67	46,06	75,4	74,6
Amor	36,5	33,25	65,5	64,7
Emocposit	38,42	41,88	54,3	56
Emocnegat	47,11	48,24	67,7	68,1
Animació	20,72	21,41	52,6	56,9
Relaxació	15,99	17,28	53,9	51,3

Taula 27. Comparativa dels valors d'efectivitat amb i sense l'efecte condició, en els models de regressió i classificació

La taula 27 mostra com l'efecte de la condició no incrementa l'efectivitat en la majoria de les prediccions, sinó que aquest està molt repartit. Veiem que succeeix en cadascun dels models:

- *Regressió*: En el model de regressió, l'efectivitat s'incrementa en les emocions "Alegria", "Tristesa", "Sorpresa", "Irritació" i "Amor"

- *Classificació*: En la classificació, l'efectivitat s'incrementa en les emocions "Despreci", "Amor" i "Relaxació" únicament.

Com hem pogut observar, el fet d'incloure la condició no ha variat en excés l'efectivitat de la predicció en cap dels model; això pot ser degut al fet que la condició s'ha determinat a partir d'una gradació numèrica que li doni pesos diferents a cadascun dels factors; malgrat això, s'ha realitzat la mateixa prova amb les condicions com a variables categòriques i tampoc ha variat els resultats, sinó que han mantingut els resultats.

Observem a continuació quins valors d'efectivitat reben les condicions si analitzem les seves dades de forma independent; per fer l'anàlisi hem desglossat els descriptors i les anotacions dels subjectes per condicions, i hem realitzat models per cadascuna de les condicions. Veiem la taula 28:

Regressio (%)	Veü	Instr	Veülnst	Mescla
Alegria	-12,68	-14,83	18,89	30,84
Tristesa	27,69	11,57	6,43	22,29
Sorpresa	37,08	39,78	2,66	2,64
Irritacio	33,81	29,12	35,19	52,3
Despreci	10,05	32,9	43,29	46,94
Amor	31,95	3,03	21,6	34,19
Emocposit	7,12	-25,77	34,28	27,95
Emocnegat	18,76	58,33	15,41	48,77
Animacio	-7,82	15,86	24,9	11,47
Relaxacio	-25,71	37,14	10,35	-13,36

Taula 28. Percentatges dels resultats de predicció per condicions en el model de regressió

La taula 28 mostra mostra quina condició obté els millors i els pitjors valors en cadascuna de les emocions, ombrejant en color verd la condició amb valors d'efectivitats més alts en aquella emoció i en color vermell l'emoció amb valors més baixos. Vegem-ho amb més detall:

- *Alegria i Emoció positiva*: obtenen els seus millors valors en la condició Mescla, i els pitjors en la condició Veu.
- *Tristesa*: obté el seu millor valor en la condició Veu, i el pitjor en la condició Veulnstrumental.
- *Sorpresa*: el millor valor l'obté en la condició Instrumental, i el pitjor en la Mescla.
- *Irritació i Amor*: obtenen els seus millors valors en la condició Mescla, i els pitjors en la condició Veu.
- *Despreci*: l'efectivitat té el seu millor valor en la Mescla, i el pitjor en la Veu.
- *Emoció negativa i Relaxació*: obtenen els millors valors en la component Instrumental, i els pitjors en la Veulnstrumental.

Veiem ara quins són els resultats en la predicció de models de classificació:

Classific (%)	Veu	Instr	Veulnst	Mescla	Total
Alegria	41,4	41,4	48,3	43,1	53,4
Tristesa	58,6	44,8	41,4	39,7	53,4
Sorpresa	41,4	34,5	50	44,8	49,6
Irritacio	60,3	74,1	72,4	77,6	70,3
Despreci	63,8	62,1	65,5	74,1	74,6
Amor	50	37,9	56,9	48,3	64,7
Emocposit	34,5	51,7	29,3	46,6	56
Emocnegat	37,9	48,3	60,3	56,9	68,1
Animacio	53,4	48,3	56,9	53,4	56,9
Relaxacio	32,8	63,8	48,3	50	51,3

Taula 29. Percentatges dels resultats de predicció per condicions en el model de classificació

La taula 29 mostra els següents resultats:

- *Alegria*: obté el seu millor valor en la condició VeulInstrumental i comparteix el valor mínim en les condicions Veu i Instrumental.
- *Tristesia*: obté el seu millor valor en la condició Veu i el pitjor en la condició Mescla.
- *Sorpresa, Desprezi, Amor i Animació*: tenen els seus millors valors en la condició VeulInstrumental i els pitjors en la condició Instrumental.
- *Irritació*: obté els seus millors valors en la condició Mescla, i els pitjors en la condició Veu.
- *Emoció positiva* : obtenen els millors valors en la component Instrumental, i els pitjors en la VeulInstrumental.
- *Emoció negativa*: els millors valors els presenta en la condició VeulInstrumental i els pitjors en la Veu.
- *Relaxació*: els millors valors es donen en la condició Instrumental i els pitjors en la Veu.

5.2.4. Anàlisi dels resultats de la combinació de condicions

L'anàlisi dels resultats de la combinació de condicions té com a objectiu analitzar dos escenaris diferents en el reconeixement d'emocions en la música; el primer pretèn comparar els resultats de la predicció de la condició VeulInstrumental amb els resultats del model combinat amb els descriptors d'aquesta condició i les condicions dels elements que la componen (Veu i Instrumental); el segon, pretèn realitzar un procediment similar amb la condició Mescla. Els resultats d'aquesta comparativa permeten observar si les prediccions amb elements d'una cançó per separat milloren i donen lloc a un model més efectiu.

Model combinat VeulInstrumental

Per realitzar el model combinat VeulInstrumental, es junten els descriptors de les condicions que componen la mescla (Veu i Instrumental) juntament amb els de la condició a comparar (VeulInstrumental) i es crea un model a partir de les anotacions de la condició VeulInstrumental.

Veiem la taula 30:

%	Regressió		Classificació	
	Combinat	Original	Combinat	Original
Alegria	37	30,84	63,8	43,1
Tristesa	31,7	22,29	60,9	39,7
Sorpresa	40,7	2,64	63,8	44,8
Irritacio	55,7	52,3	81	77,6
Despreci	54,8	46,94	78,2	74,1
Amor	40,06	34,19	67,2	48,3
Emocposit	43,9	27,95	59,2	46,6
Emocnegat	43,03	48,77	63,2	56,9
Animacio	48,4	11,47	56,9	53,4
Relaxacio	38,4	-13,36	59,8	50

Taula 30. Comparativa dels resultats de regressió i classificació del model combinat amb Veu, Instrumental i Veu Instrumental i el model de la condició VeulInstrumental

La taula 30 mostra com en la majoria de les emocions, l'ús del model combinat VeulInstrumental millora els resultats d'efectivitat tant dels models de regressió (excepte en el cas de l'emoció negativa) com en els models de classificació.

Model combinat Mescla

Per realitzar el model combinat Mescla, es junten els descriptors de les condicions que componen la mescla (Veu i Instrumental) juntament amb els de la condició a comparar (Mescla) i es crea un model a partir de les anotacions de la condició Mescla.

Veiem la taula 31:

%	Regressió		Classificació	
	Original	Combinat	Original	Combinat
Alegria	18,89	30	48,3	56,3
Tristesa	6,43	38,7	41,4	63,8
Sorpresa	2,66	21	50	52,9
Irritacio	35,19	57,7	72,4	71,83
Despreci	43,29	43,2	65,5	67,8
Amor	21,6	36,6	56,9	55,2
Emocposit	34,28	44,6	29,3	55,2
Emocnegat	15,41	44,8	60,3	63,8
Animacio	24,9	30,5	56,9	48,3
Relaxacio	10,35	17,5	48,3	49,4

Taula 31. Comparativa dels resultats de regressió i classificació del model combinat amb Veu, Instrumental i Mescla i el model de la condició Mescla

La taula 31 mostra com en la majoria de les emocions, l'ús del model combinat Mescla millora els resultats d'efectivitat tant dels models de regressió (excepte en el cas de l'emoció Despreci) com en els models de classificació (excepte en les emocions Irritació, Amor i Emoció Negativa).

6.- Discussió

L'anàlisi dels resultats al llarg del projecte indica, de forma general, un efecte rellevant de les condicions en el reconeixement de les emocions, que podem veure en les dades obtingudes tant de l'anàlisi descriptiu com de la comparativa dels resultats del modelatge. Aquest efecte no és el mateix per totes les emocions, sinó que cadascuna d'elles presenta uns pesos diferents en funció de la condició seleccionada. Si observem amb detall les dades és difícil establir una tendència per cada emoció en el procediment de reconeixement d'emocions, ja que els resultats del modelatge d'emocions no reforcen les hipòtesis creades a partir de l'anàlisi descriptiu d'aquestes. Veiem quins són els comportaments observats per cadascuna d'elles:

- *Alegria*. Tendeix a ser una emoció amb un grau de percepció mig, concentrat principalment en la condició Mescla i amb poc pes en la condició Veu, tal i com s'observa i coincideix amb els resultats obtinguts en la creació de models de predicció; els resultats de la predicció s'han vist accentuats en els models combinats, pel que es manté la posició de rellevància de l'efecte condició.

- *Tristesa*. De la mateixa forma que amb Alegria, Tristesa és una emoció amb un grau de percepció mig, concentrat especialment en la condició Veulinstrumental però amb pesos força repartits. En el modelatge presenta uns resultats força similars, que coincideixen en major mesura amb el model de classificació, el qual presenta uns resultats de predicció en la Veu alts i més baixos en la mescla. En els models combinats s'han vist incrementats els resultats de la predicció, pel que es manté la posició de rellevància de l'efecte condició.

- *Sorpresa*. Amb un grau de percepció en la posició mitja-baixa, els seus graus de percepció més alts es concentren en la part instrumental, amb una distribució força uniforme en la resta de condicions; els resultats del modelatge reforcen els pesos descrits, principalment en el model de regressió, en el qual la condició Instrumental presenta els valors més alts de predicció. En els models combinats s'han vist incrementats els seus resultats, pel que es manté la posició de rellevància de l'efecte

condició.

- *Irritació*. De la mateixa forma que l'emoció Sorpresa, Irritació presenta uns valors de percepció mitjos-baixos; el valors alts es concentren en la condició Veu i els baixos en la Instrumental. Els resultats del modelatge es mostren força incoherents, i en cap dels models la condició Veu és la que obté els valors més alts; els models combinats han vist reforçat o incrementat el seu grau de percepció excepte en el cas del model de classificació per al combinat de Mescla.

- *Despreci*. Amb un grau de percepció molt baix, el pes dels resultats recau en la Veu i molt poc en la part instrumental; de la mateixa forma que amb l'emoció Irritació, els resultats del modelatge es mostren força incoherents, malgrat que en els models combinats s'han vist incrementats els resultats de la predicció.

- *Amor*. Tendeix a ser una emoció amb un grau de percepció mig, concentrat principalment en la condició Veu i amb poc pes en la condició Instrumental. Els resultats del modelatge es mostren força coherents, ja que la condició Instrumental es presenta en els dos models amb els valors més baixos de predicció; la majoria de models combinats (excepte el model de classificació en el combinat de Mescla) han vist incrementats els seus resultats, pel que es manté la posició de rellevància de l'efecte condició.

- *Emoció positiva*. Amb un grau de percepció força mig, concentra els seus resultats de predicció en la condició Mescla, amb molt poc pes en la veu. Els resultats de la classificació són força incoherents, malgrat que en els models combinats s'ha vist incrementada la seva percepció.

- *Emoció negativa*. Presenta un grau de percepció mig-baix, i malgrat que el seu pes es concentra en la condició instrumental, presenta una distribució entre condicions força similar; en el modelatge presenta uns resultats força incoherents, malgrat que en els models combinats s'ha vist incrementada la seva percepció (excepte en la regressió del model VeuInstrumental).

- *Animació*. El grau de percepció és força alt i els seus pesos es concentren principalment en la condició Mescla; els resultats del modelatge es mostren força

coherents, malgrat que no coincideixen amb els resultats de l'anàlisi descriptiu. Tot i així, en els models combinats es veu incrementada la seva percepció (excepte en la classificació del model Mescla).

- *Relaxació*. Amb un grau de percepció força baix, els valors més alts de percepció es troben en la condició Veu, i els més baixos en la part instrumental; els resultats del modelatge es mostren força incoherents, malgrat que en els models combinats es veu incrementada la seva percepció.

Les descripcions anteriors mostren emocions amb resultats incoherents entre l'anàlisi descriptiu i la creació de models; de fet, malgrat que els resultats de l'anàlisi descriptiu mostraven amb claredat que la condició tenia un efecte rellevant sobre la percepció d'emocions, quan s'ha realitzat la creació de models amb la condició com a descriptor predictor, els valors de predicció no han variat pràcticament; això pot ser degut a diversos motius,: que la tècnica d'aprenentatge automàtic li donés pesos valors diferents als valors de la condició (1,2,3,4), que el programa utilitzat no entengués la variable condició de forma categòrica en el moment que s'han categoritzat les condicions o que els descriptors utilitzats no fossin els adients per poder observar aquesta rellevància.

Si aquest efecte l'observem en detall en cadascuna de les emocions, com hem fet en l'anàlisi de tendències emoció a emoció, podem veure com la majoria d'emocions presenten resultats força incoherents; en aquest sentit, es pot observar com les emocions que han obtingut uns valors més baixos en l'anàlisi descriptiu (irritació i despreci), en la primera prova de modelatge sense diferenciar condicions han obtingut els millors valors de predicció; això planteja varies hipòtesis: d'una banda és possible que malgrat que aquestes emocions mostren valors més baixos en l'anàlisi descriptiu, quan es mostren en un fragment d'àudio es mostrin de forma molt clara, i per tant els resultats entre subjectes siguin força coherents, i això fa que els resultats de la predicció s'incrementin; d'una altra banda pot passar que aquestes emocions, en les que el pes recau en la condició veu, estiguin correlacionant força bé amb els seus descriptors, tot i que aquesta suposició no és massa vinculant ja que en altres emocions amb el pes en la condició veu no ha succeït.

Respecte a les emocions que han donat valors força coherents, crida l'atenció que en la majoria d'elles la distribució dels pesos està força repartida entre condicions; de fet,

si observem l'anàlisi descriptiu, veurem com Tristesa i Sorpresa no mostren un efecte rellevant de la condició, amb el que pot indicar que el fet d'haver realitzat un anàlisi amb els descriptors i les anotacions de totes les condicions de forma global, els valors de predicció es vegin decrementats.

Respecte als models utilitzats en la recerca, podem dir que tot i el fet que el model classificador fa una predicció categoritzada i per tant es redueix la precisió dels valors a prediure, aquest obté uns resultats de predicció més alts que en els de regressió en la majoria dels casos; no obstant, la majoria de modelatges han mostrat una efectivitat del 50% que no és més que la probabilitat que es té quan llençes una moneda a l'aire; això fa pensar en que tant la metodologia utilitzada com el material utilitzat no hagin estat els més adients; en aquest sentit, l'experiència obtinguda al llarg del procés de desenvolupament del projecte em fa pensar que si s'hagués de repetir tot el procediment, pendria unes altres decisions al respecte d'alguns aspectes, que detallo a continuació:

- *Selecció de descriptors.* Els descriptors són un dels aspectes a millorar en futures investigacions, d'una banda per què tres quartes parts d'aquests són descriptors sorgits a partir de l'anàlisi musical, mentre que únicament una quarta part està destinada a l'anàlisi de la veu; aquesta descompensació ha pogut fer que els descriptors de la veu no fossin rellevants; a més de les tres parts de descriptors de l'anàlisi musical, una gran part són descriptors envers la distribució d'energia en l'espectre, mentre que no s'ha disposat de suficients descriptors de tonalitat. D'una altra banda gran part de les mesures de l'anàlisi s'han realitzat a partir de descriptors de baix nivell, trobant a faltar descriptors que per la seva dificultat d'anàlisi haguessin pogut donar una informació valuosa envers la percepció d'emocions, com són el mode, el vibrato o la intervàlica.

- *Qualitat dels intèrprets en els enregistraments.* El fet d'haver seleccionat gravacions de músics aficionats i semi-professionals ha pogut fer minvar els resultats de predicció, degut a que cap la possibilitat que aquests intèrprets no fossin tant hàbils en la comunicació de les emocions, i per tant els oients no poguessin percebre les emocions clarament.

- *Selecció de les emocions.* Malgrat que la selecció s'ha realitzat cobrint totes les dimensions de valència i activitat possibles, algunes de les emocions escollides han

estat probablement difícils d'avaluar degut al poc ús i la interpretació que s'hagi fet d'aquesta en cadascun dels subjectes.

7.- Conclusions

Al principi del projecte es van plantejar el proporcionar una visió general de les característiques i problemàtiques més destacades en la classificació d'emocions expressades en la música, així com de l'efecte de les parts d'una cançó (veu, instrumental i producció) en la percepció de les emocions..

Com a resultat, al present projecte podem concloure doncs que les diferents parts d'una cançó (part vocal, part instrumental, producció), tenen un efecte rellevant en la predicció dels models de reconeixement d'emocions incrementant els seus resultats en la majoria dels casos.

D'una altra banda, al comparar els resultats dels models de regressió i classificació utilitzats, hem pogut observar com el model de classificació obté uns millors resultats en la predicció, a detriment d'oferir una predicció categòrica amb tres respostes possibles.

Finalment, els resultats obtinguts a partir de l'anàlisi exhaustiu de les emocions ha demostrat en alguns casos que encara hi ha molt treball a fer en el camp de la ciència afectiva i la psicologia musical, ja que la complexitat de les emocions i els seus mecanismes fan que l'avaluació i l'anàlisi dels resultats presentin una complexitat amb la qual és difícil obtenir unes tendències i uns patrons de comportament clars.

7.1.- Contribucions

El present projecte ha servit com a punt de partida per l'anàlisi de les emocions des d'un punt de vista diferent, centrat en descomposar els elements d'una cançó en diverses parts de forma que cadascuna d'elles tingui una entitat pròpia que ajudi a desenvolupar models de classificació més eficients. A partir d'aquest desenvolupament s'han pogut observar característiques de les emocions en cadascuna de les condicions, així com les tendències més rellevants de les emocions en cadascuna d'elles; a més, en la segona part del projecte, centrada en el modelatge d'emocions,

s'han realitzar comparatives de models i s'ha pogut donar informació envers el comportament dels models en cadascuna de les emocions i condicions.

Finalment, el desenvolupament del projecte ha acumulat una sèrie d'arxius que poden ser utilitzats tant per posteriors recerques amb elements d'una cançó separats, com per automatitzar processos d'anàlisi o questionaris mitjançant els scripts creats, com per realitzar anàlisis descriptives o models predictius a partir de les dades dels subjectes i mesures obtingudes.

Així doncs, crec que el present projecte és una bona eina de suport que, tant pel recull d'informació realitzat envers aquest àmbit, com pels resultats i materials obtinguts a partir del desenvolupament del projecte, espero proporcioni un punt de partida en futures recerques.

8.- Referències

BOGDANOV, D., WACK N., GÓMEZ, E., GULATI, S , HERRERA. P., MAYOR, O., et al. (2013). *Essentia, an Audio Analysis Library for Music Information Retrieval*. International Society for Music Information Retrieval Conference.

CACIOPPO, J. T.; GARDNER, W (1999) “*Emotion*”. *Annual Review Psychology*. 50:191-214

CARTWRIGHT, M.; PARDO, B; REISS, J (2014). *Mixploration: Rethinking the Audio Mixer Interface*. Queen Mary University of London.

EKMAN, P (1974). *Darwin and Facial Expression*. Academic Press.

EYBEN, F.; SALOMAO, G.; SUNDBERG, J.; SHCERER, K.; SCHULLER, B.W. (2015). “Emotion in the singing voice -a deeper look at acoustic features in the light of automatic classification”. *Eurasip Journal on Audio, Speech and Music Processing*. 2015:19

FEHR, B.; RUSSELL, J. (1984) “*Concept of emotion viewed from a prototype perspective*”. *Journal of Experimental Psychology General*. 113(3) :464-486

FLETCHER, M. (2011). *The Effect of Spatial Treatment on Listener’s Emotional*. Asarp. Disponible a:

<http://arpjournal.com/the-effect-of-spatial-treatment-of-music-on-listener%E2%80%99s-emotional-arousal/>

GRABAN, E. (2015). Subjective evaluation of the audio bandwidth extension program MaxBass. Norwegian University of Science and Technology. Disponible a: https://brage.bibsys.no/xmlui/bitstream/handle/11250/2371553/13915_FULLTEXT.pdf?sequence=1&isAllowed=y

GOLEMAN, D. (1995). *Emotional intelligence*. New York: Bantam publications

JUSLIN, P.; SLOBODA, J. (2010). *Handbook of music and Emotion*. New York: Oxford University Press.

KLEINGINNA, JR; KLEINGINNA, ANNE M. (1981). “*Motivation and Emotion*”. 5: 263

MOFFAT, D; RONAN, D.; REISS, J. D. (2015). *An Evaluation of Audio Feature Extraction*

Toolboxes. Center for Digital Music. Queen Mary University of London. Disponible a:
https://www.ntnu.edu/documents/1001201110/1266017954/DAFx-15_submission_43_v2.pdf

PLUTCHIK (1994). *The Psychology and Biology of Emotion*. HarperCollinsCollege publications
ROSCH, E (1978). "Principles of categorization". University of California, Berkeley

RONAN, D; REISS, J.D; GUNES, H. (2016). *The Relationship Between Emotion and Music Production Quality*. Queen Mary University of London. Disponible a:
<https://www.eecs.qmul.ac.uk/~josh/documents/2016/Ronan%20DMRN%202016.pdf>

WILSON, A.; FAZENDA, B.: (2016). *Variation in Multitrack Mixes: Analysis of Low-Level Audio Features*. University of Salford.

WITTEN, I.H.; FRANK, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd edition. San Francisco.