

# Дипломна работа

Диана Генева <dageneva@qtrp.org>

2019

# Съдържание

1	Нулева зона	2
2	Емоции	3
3	Сигнал от реч	7
3.1	Физика на тъгата . . . . .	7

# Глава 1

## Нулева зона

Бла, бла, бла, аз съм толкова емоционална. Не знам



# Глава 2

## Емоции

За да кажем как ще разпознаваме емоции в сигнали, трябва да изберем две неща. Първо, какво ще наричаме емоция и второ - кои емоции ще класифицираме.

Най-точното описание, което може да се даде за емоция, е може би цитатът на Потър Стюърд (член на Върховния съд на САЩ): “Когато го видя, ще го разпозная”<sup>1</sup>. Изкушаващо е да повторя думите му: “Няма днес да се опитвам точно да дефинирам материята, попадаща под това кратко описание; възможно е никога да не мога да дам разбираемо описание. Но едно нещо знам, когато го видя, ще го разпозная.”. Но за съжаление, за да можем да продължим напред в текста на дипломната работа, ще трябва да разботим с малко по-конкретни дефиниции. Един от най-ранните опити да се дефинира емоция, идва от Чарлз Дарвин и книгата му “За изразяването на емоциите при човека и животните” от 1872 г. Той използва емоцията, за да потвърди еволюционната си теория, тъй като според неговите наблюдения изразяването на емоциите е сходно между животните и хората, и се обяснява еволюционно. В книгата си, Дарвин представя три принципа, които описват изразяването на емоциите:

### 1. “Принцип на полезните навици”

Такива са смръщване на вежди, при което в очите влиза по-малко светлина, или вдигане на вежди, което увеличава полето на зрение. Когато човек е учуден, той вдига вежди, за да “види по-ясно ситуацията”. Такива навици, според Дарви, се предават наследствено заради полезния си характер.

### 2. “Принцип на противоположностите”

Това са навици, които нямат практическа полза, но представляват противоположност на някакъв друг естествен навик. Пример за това е свиването на рамене, което е израз противоположен на уверено или агресивно изразяване.

### 3. “Принцип на нервните сигнали”

---

<sup>1</sup>I know it when I see it

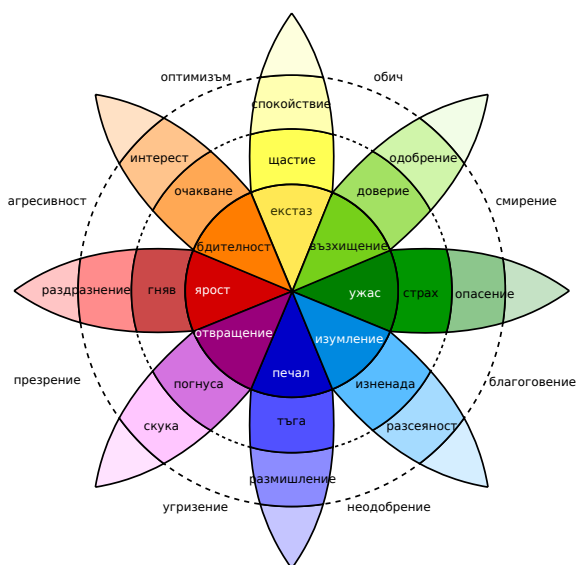
Когато се говори за Даривновата теория, най-често се има предвид този принцип. Според него, някои навици са породени от сигнал от нервната система. Такива са например изявите на гняв и страх.

Тъй като тези изрази на емоция се предават еволюционно, то те се наблюдават у всички животни (в частност и хора).

Този представа за емоциите е продължена от Робърт Плутчик (1927 - 2006). Базирайки се на еволюционната теория, той се опитва да опише понятието "емоция". Наблюденията му са следните:

Еволюционния произход на някои емоции е по-явен от при други. Например изразът на страх при животните и хората е много сходен и цели едно и също - да премахне причинителя. Изразът на любов обикновено е с репродукционни цели. Някои емоции обаче са по-сложни и първоизточникът им се описва по-трудно, затова търсим начин да генерализираме понятието. Емоцията е сложна верига от събития, която започва с някакъв стимул. В следствие настъпва фаза на "изпитване на емоция" и фаза на физиологични промени (тук възниква научният спор за кокошката и яйцето - дали чувството е първо или физиологичните промени). Те предизвикват целенасочено гържане, което цели да премахне дразненето на стимула и да върне състоянието на еквилибриум.

Плутчик прави психо-еволюционна класификация на емоциите, като избира осем главни емоции - две по две противоположни. Всяка от тези базови емоции е свързана с поведение важно за оцеляването. Например страх и поведението "бий се или бягай", свързано с освобождаването на хормони, подготвящи организма за бягство или битка. Всяка от не-базовите емоции се получава като комбинация на базовите. Плутчик сравнява това с комбинирането на цветове, което може да се види и на известното му "Колело на емоциите" (1980), показано на [Фигура 2.0.1](#).



Фигура 2.0.1: Колелото на емоциите на Плутчик

Базовите емоции са: страх, изненада, тъга, познуса, гняв, очакване, щастие, доверие. По “листенцата” на всяка от тях се намират по-силният и съответно по-слабият вариант. За съжаление, колкото повече класове искаме да класифицираме, толкова повече данни са нужни, което прави използването на класификацията на Плутчик трудно за практически цели. Друг проблем е, че различните емоции имат различна продължителност на проявлението, като например тъгата може да продължава с дни, докато изпитването на познуса е много по-моментно. Това прави съставянето на база данни за емоционално разпознаване трудно.

След като вече започнахме с цитат, време е и за другото клише, а именно етимология. Думата емоция произлиза от френското *émouvoir* (вълнувам, възбуждам интерес) и по-назад латинското *emoveō* (местя навън/извън пътя). Затова е и естествена класификацията на база как “вълнува” съответната емоция. Известна такава класификация е PAD моделът, разработен от Алберт Мейерабиан и Джеймс Ръсел (1976), в която всяка емоция се описва в тримерно пространство с оси “удоволствие”, “ниво на възбуда”, “доминантност”<sup>2</sup>. Оста за удоволствие определя дали емоцията е приятна или не, оста за нивото на възбуда описва колко енергия е нужна за изразяването на емоцията, а доминантност - колко доминиращ или съответно покорен се чувстваш под влиянието на емоцията. По-често се използва опростен модел наречен валентност-активация, който използва само две от осите (удоволствие и ниво на възбуда). Изразяването на някои емоции в този модел е показано на [Фигура 2.0.2](#)

За да изберем какъв модел за класификация на емоциите да ползваме, нека разгледаме какви са силните и слабите страни на двете изследвани области. Според [survey], в речта лесно се разпознават характеристики свързани с активацията, тъй като нивото на активацията се отразява на енергия на говора. Тоест емоции като щастие и гняв (които са с висока активация) карат хората да говорят много по-силно и възбудено, докато емоции като тъга и умора (с ниска активация) са свързани с много по-пасивно поведение като цяло. Дадените примери за разпознаване на емоциите в мозъка в [brain\_survey], показват, че по-лесно се разпознава валентност, макар че може да бъде извлечена информация и за активацията.

При изборът на емоции, които да изследваме, трябва да изберем такива, за

<sup>2</sup>Pleasure, Arousal, Dominance



Фигура 2.0.2: Модел “валентност-активация”

класификацията на които ще помогнат и двата сигнала. В случая ще разглеждаме четири емоции - три от които са част от базовите емоции на Плутчик, а четвъртата емоция е “неутрална емоция”. Изборът е такъв, че хем да нямаме прекалено много класове, хем тези емоции да могат да се предизвикват лесно в опитна среда и да могат да се описват вербално. Финално избраните емоции, които ще се опитваме да разпознаваме в сигнали от реч и ЕЕГ, са следните:

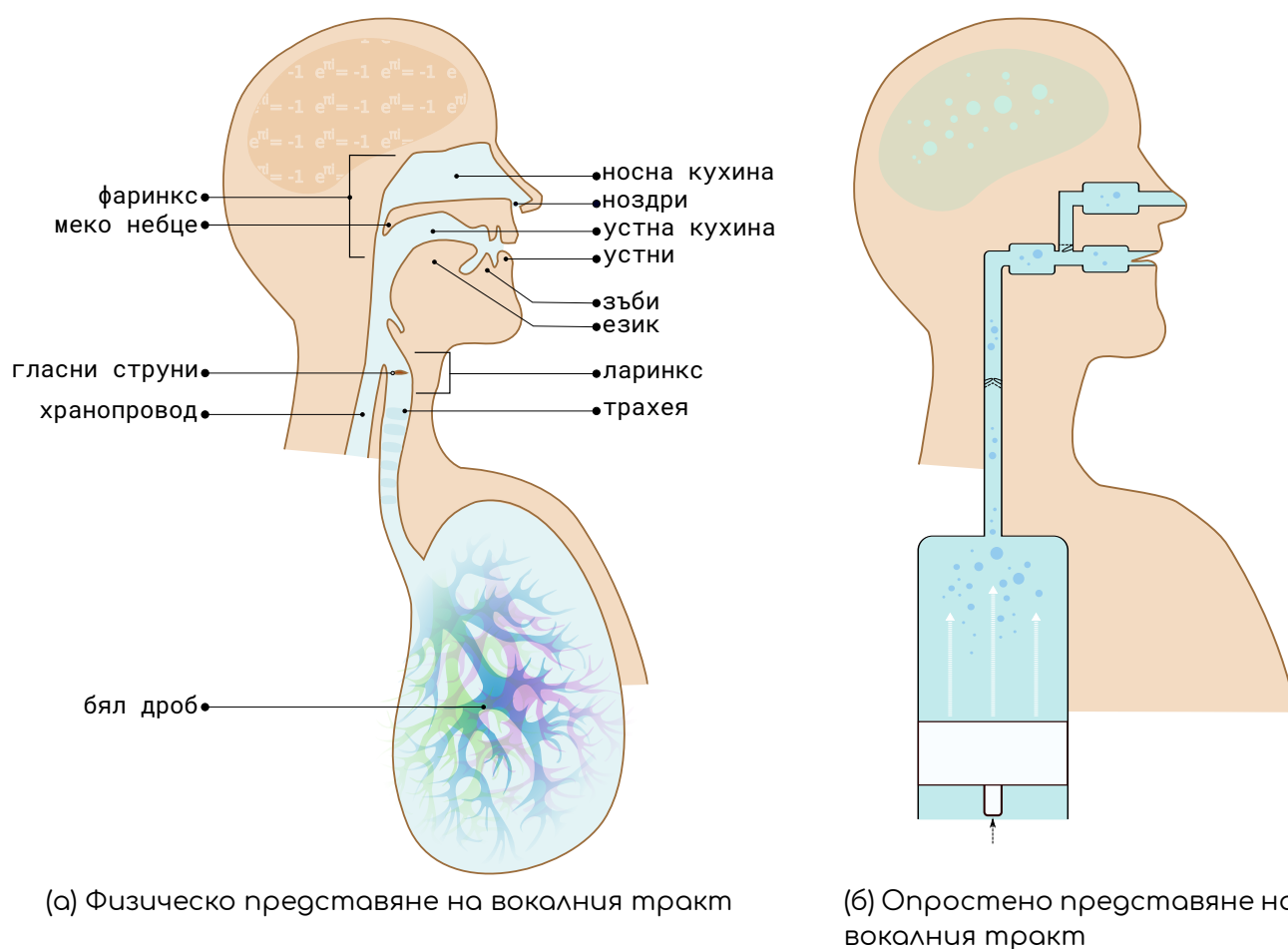
1. **Гняв** - висока активация, отрицателна валентност
2. **Щастие** - висока активация, положителна валентност
3. **Неутрална емоция** - неутрална активация, неутрална валентност
4. **Тъга** - ниска активация, отрицателна валентност

След този бегъл опит да дефинираме понятието емоция и да изберем кои класове емоции ще изследваме, трябва да разгледаме свойствата на двата входни сигнала. Нека започнем от сигналите от реч.

# Глава 3

## Сигнал от реч

### 3.1 Физика на тъгата



Фигура 3.1.1: Система за производство на реч

Вокален тракт е общото название на кухините над ларинкса (гръкляна), през които минава въздухът при произвеждане на реч. При хората той се състои от ларингеална кухина (съдържаща ларинкса и гласните струни), фаринкс, устна кухина и носна кухина, както може да се види на Фигура 3.1.1а. Вокалният тракт е отговорен за генериране на различни звуци, като текущата конфигурация на отделните му компоненти определя какъв ще бъде сами-



ят звук. Снорег [emotional:shit], освен от вида на този звук, конфигурацията на вокалния тракт зависи и от емоцията, която изпитва говорещият. Смята се, че емоционалното състояние е пряко свързано с определени промени в организма, например ускорено дишане или мускулно напрежение, а тези промени се отразяват върху произведената реч. Често дори ефектите от тези промени са станали нарицателно за самата емоция. Из българската литература се срещат изречения като „страхът стискаше гърлото, задушаваше гласа“ [talev], а изрази като „буца в гърлото“ или „пресъхнало гърло“ са навлезли в разговорната реч като асоциации на „тъга“. Тъй като изпитваната емоция влияе пряко на конфигурацията на вокалния тракт, бихме искали да извличаме характеристики, които описват тази конфигурация.

Да разгледаме по-подробно Фигура 3.1.16, която описва цялостна система за производство на реч в по-опростен вариант. Речта, всъщност, представлява просто акустичната вълна, получена на края на системата - устни и ноздри - в следствие на изтласквания от белия дроб въздух.

Белият дроб работи като енергиен източник за тази системата - въздушният поток, получен при свиването му от междуребрните мускули и диафрагмата, се пропъгува нагоре по трахеята и през глотиса (отвора между гласните струни).

Действието на глотиса може да се види най-ясно при произнасяне на гласна. Гласните струни пропускат пропъгувания въздух. Тъй като глотисът е стеснение, налягането в него в този момент е по-малко от това в който и да е от двата му края. Съгласно закона на Бернули в някакъв момент то става толкова ниско, че позволява на гласните струни да се затворят. В следствие се натрупва налягане зад гласните струни заради тласкания от белия дроб въздух, което в някакъв момент ги принуждава да се отворят, и цикълът се повтаря отначало. В резултат се получава осцилиране на гласните струни. Честотата на отварянето и затварянето зависи от анатомични особености като еластичността и големината на гласните струни, налягането в белия дроб и други.

При мъжете тази честота е средно 125 Hz, а при жените - 210 Hz.

Акустичната вълна, която се получава в следствие на осцилацията, преминава през вокалния тракт, където се завихря при срещане на прегради като устни и зъби и в крайна сметка напуска системата през някой от отворите.

При този процес се губи част от енергията поради различни фактори като съпротивлението на въздуха и поглъщането на вълната от меките и еластични стени на вокалния тракт.

В зависимост от начина, по който вълната напуска системата, можем да класифицираме произведените звуци по следния начин:

1. Озвучени

При тези звуци гласните струни осцилират квази-периодично.

2. Проходни (фрикативни)

При образуването на проходни звуци, вълната среща преграда по пътя

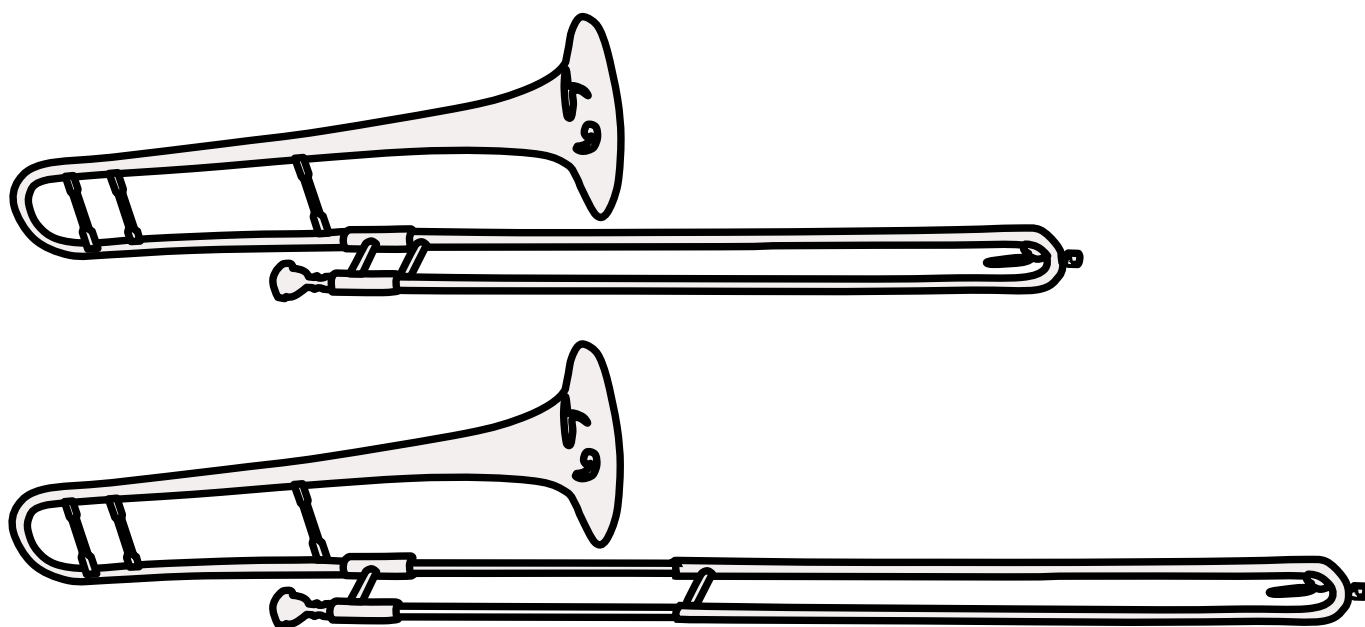
си (като например зъби, устни) и се получава турбуленция, при опита да бъде избутан въздухът през презградата.

### 3. Презградни (експлозивни)

Те се получават при напълно затворена презградата, за която се натрупа налягане, което се освобождава рязко чрез отваряне на презградата.

Обикновено речта, която произнасяме, е разделена на думи, като отделните звукови единици в тях се наричат фонemi. За да се произнесе определена дума, вокалният тракт трябва да застане в правилната конфигурация за следващата фонема в думата. Когато вокалният тракт се наглася за дадена фонема, настъпват промени, като например стените на устната кухина се приближават или мекото небце, служещо като клапа към носната кухина, се затваря. Може да се усети, че при изговаряне на „а“ отворът е много по-голям, отколкото при произнасяне на „у“. Тази промяна влияе върху спектралните свойства на вокалният тракт.

Нека за улеснение си представим, че сме моделирали вокалният тракт с последователност от тръби, за да се абстрахираме от сложната му физическа структура. Тогава при смяна на фонемата, се променят дължината и диаметърът на тръбите. Това влияе на времето, за което акустичната вълна ще стигне до края на тръбата и съответно на честотата, на която ще се получи резонанс. Тоест в зависимост от конфигурацията, ще се усилят или затихнат различни честоти, спрямо резонанса. Това свойство се нарича честотна пропускливост. Идеята лесно се вижда при свиренето на духови инструменти.



Фигура 3.1.2: Тромбон

При тях по някакъв начин се променя изходът на вълната, например отпушване и запушване на дупки, и съответно честотата, на която се получава резонанс, тъй като пътят на вълната е скъсен или удължен. Както може да се види на [Фигура 3.1.2](#), при тромбона буквално се сменя дължината на тръбата, което означава, че на вълната ѝ трябва повече време, за да се отрази,

тоест резонансът е на по-малка честота и съответно изходящият звук е по-нисък.

Това значи, че ако знаем как се пропазира вълната по отделните тръби на вокалния тракт и какви са спектралните свойства накрая, можем да съдим за текущата му конфигурация. В такъв случай, за да изследваме подлежащата емоция при реч е нужно да изследваме тези свойства в достатъчно кратък отрязък от време, в който конфигурацията е статична. Обикновено се приема, че този период е между 10 и 20 милисекунди ([rabiner\_schafer78]).

В следващия раздел ще разгледаме как можем да моделираме вокалния тракт с модела на тръбите, за да можем да извлечем спектралните му свойства.