Language as a Linguistic Art:
The Expression of Emotion in the
Spoken and Written Word

Matthew D Coombes

MA Thesis - Linguistic Aesthetics

De Montfort University

- November 2014 -

Introduction

Millions of years ago, man would have employed the use of language to instigate an emotional response in the brain of his fellows, whilst today the role of language has shifted to satisfying our instinct for knowledge. It is therefore accepted that language and voice have continued to separate from the emotional centres of the brain during the course of human evolution (Perlovsky, 1999). However, connections still exist between the limbic (emotional) centres and the cortices of the brain, meaning that influence between both remains possible. Consequently, the field of aesthetics has become the subject of interdisciplinary research, with the shared purpose of ascertaining the stimuli responsible for the encoding of emotion in language and speech. From those areas noted to the present day, the collective research can be categorised into three areas of study, of which the present paper will focus on phonaesthetics, whilst taking influence from the other areas.

1. **Neurolinguistics** - an area closely linked with psycholinguistics and concerned with the study of the parts of the brain responsible for the acquisition, production and comprehension of language.

2. **Phonaesthetics** - the study of the inherent valence and arousal of language, and the paralinguistic features attributed to the phenomenon.

3. **Social Aesthetics** - the study of the social, pragmatic and behavioural aspects of communication, and the means by which they can attribute to the emotional encoding of communication.

Since Humboldt identified human language as a system governed by rules as opposed to a set of words with associated definitions, the means by which language is interpreted took on a new meaning. Determined that a connection existed between the sound of a word and its meaning, Humboldt stated that rather than being simply imitative of the definition they represent, the sounds of a language were the "very fabric of thought" (Pütz, 2000); the thoughts

themselves being reproduced as an internal dialog, using the grammar and vocabulary of the individuals native language.

This model, published in 1836, was a response to the accepted notion of language being a process of naming which, based on a finite group of predetermined ideas, assumes cognition and language to be separate abilities of the individual. Conversely, Humboldt's approach identifies the way we speak as influencing our thought processes, or what later became referred to as the *linguistic relativity principle* [1].

In support of this idea, Saussure argued that should language be based on predefined concepts, then all languages would share exact equivalents in their vocabulary. Knowing this not to be true, he defined language as a product which is "passively assimilated by the individual" (Saussure 1916:14), whereby *sound-images* are associated with concepts. The sound-image is a term initially used by Saussure, not to describe a material sound, but the psychological imprint that the sound has in our brain. Therefore, whilst any given concept can be considered universal, the sound-image it represents remains unique to the language in which it is spoken, as the diagram below illustrates.

Figure 1 - Saussure's concept of sound-images and concepts

---

[1] The linguistic relativity principle is better known as the *Sapir-Whorf Hypothesis*, whilst named as a result of the combined works of Edward Sapir and Benjamin Whorf, was predicated on the works of Humboldt and Saussure, amongst others.

Sapir took the characterisation of sounds further having shown that the acoustic qualities of specific phonemes were regularly associated with a similar signified object [2]. This conclusively proved that meaning can be taken directly from the sound, or phoneme, regardless of their applied significance, and that vocalic and consonantal contrast had consistent symbolic significance, of which had little relation to the associative values of the words. The way in which phonemes were put together, in what Sapir referred to as "phonetically possible" or "phonetically impossible" [3], also showed little difference to the results (Sapir, 1963:63). Today this field of study is commonly referred to as sound symbolism, or *phonosmbolism*, and is a continued area of linguistic research.

Fónagy then continued the results gleaned by Sapir's experiments, the results of which compiled in *Die Meatphern in der Phonetik* (tr. *The Metaphors in the Phonetics),* and published in 1963. His work correlated phonemes with metaphors, somewhat subjectively, using European languages. For example, front vowels (e, i, a) are considered "light" and "high", whereas back vowels (ɑ, o, u) are "dark" and "heavy". According to Fónagy, the former can elevate positive emotions, such as happiness and peace, with the latter signifying the contrary. The encoding of such metaphoric data is attributed to paralinguistic phenomena such as pitch, intonation and amplitude variation. These messages are then decoded by the auditor in a preverbal semiotic system responsible for aesthetic and emotional attitudes (Fónagy 1999:3, Podhorodecka, 1999:3). Observations that indicated a promotion of positive emotion also showed that the relative duration of vowel articulation is increased, whilst plosive articulation is decreased (Fónagy, 1991: 495-496). The diagram below interprets this model (figure 2).

---

[2] Sapir's experiments conducted in the 1930's consisted of a number of word pairs - half of which being unfamiliar to the subject, thus removing the influence of meaningful linguistic association. Words such as "mil" and "mal" would show the former to be associated with smaller references, whilst the latter to signify larger objects. Such results showed a clear correspondence between vowel backness and an objects relative size.

[3] In Sapir's work, the term "phonetically possible" referred to words which demonstrate sequences that frequently appear in the English language, whilst "phonetically impossible" signified the opposite, or in Saussurean terms, a sound-image with no associated concept.
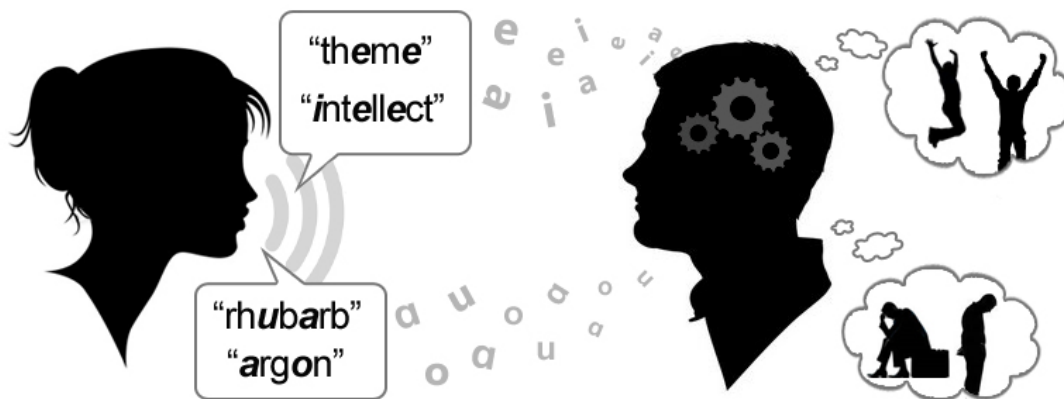
Figure 2 - Fónagy's model of correlating phonemes and metaphors:
front vowels promote positive emotion, whilst back vowels
promote negative emotion

Despite the assertion that language has continued to separate from our emotional responses, as evolution has preserved the connections between the two, we have become subject to an advancing state of neurological development whilst maintaining the primordial influences. However, the correct level of cognitive and emotional engagement is crucial for developing language models. Should a language remain highly emotional (as it was among primates), such models would have no means to develop. Conversely, should there be no emotion at all, there would be no motivation to engage in conversation, thus diminishing the likelihood of developing higher cognitive models (Perlovsky, 2009:2). Therefore, as we understand that emotion exists in language, it is possible to appropriate individual communicative characteristics in order to stimulate specific areas of the brain. In a majority of cases, the alteration of phonetic characteristics such as such as voice, pitch or intonation, are employed for specific purposes [4], and by ascertaining which emotion pertains to a given phoneme, it is possible to assign an "emotional score" to any given dialogue, based on a scale derived from the specific phonemic criteria. Further experiments have been conducted to further evidence the validity of Fónagy's results, and one such case concludes with a

---

[4] Such examples may include aftershave commercials utilising deep, resonant male voice to infer masculinity, or a slow, soft voiceover to trigger empathy for a charitable advertisement.

means by which a subset of English phonemes can be measured in terms of emotionality, or what the Whissell refers to as an "emotional compass" (figure 3). What the author has again addressed, is the notion that sounds symbolise emotion, and further hypothesised that when applied to language that our understanding of linguistic communication will increase.
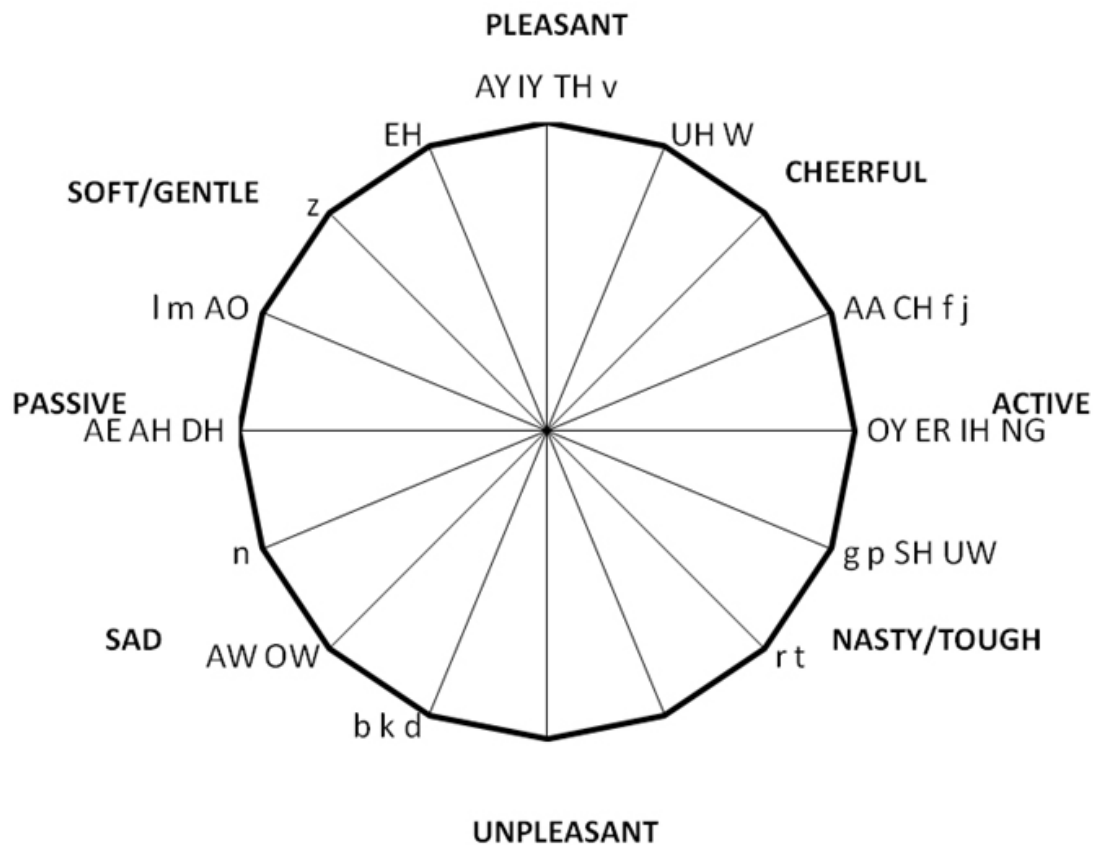

Figure 3 - Whissell's algorithm for measuring emotionality in language, or what is referred to as the *emotional compass*

Assuming that emotions can be portrayed through the phonemic aspects of language, the question remains as to what benefits can be expected in using the appropriate phonemes to stimulate emotional response. Furthermore, as a majority of studies in this area have concentrated on English as the default language by which to measure perceived emotionality, is there further support for this model through the engagement of second language speakers,

particularly those of whom have adopted sequential or simultaneous bilingualism? To what extent, if any, does the emotionality of a word take precedence over the semantic context? And can we assume cross-cultural consistency in the combined work of Fónagy and Whissell?

Whilst this addresses the native or competent language speaker, the further line of questioning occurs when we consider the same model in context of those ignorant of the semantic content. For example, given that comparatively less cognitive function is employed in the comprehension of a foreign language, can a relationship be observed between linguistic familiarity and its perceived emotionality? That is to say, does the semantically ignorant auditor experience a greater emotional concentration compared to the native speaker?

Finally, how does the same model result when considered in the context of artistic or constructed languages? Given that the variables are similar to those above, can we expect similar observations as a result?

Pidgin to Poetry: The Measuring of Linguistic Emotion

A majority of linguistic research concentrates on the conceptual exchange of information, and as a result the emotional aspects of language are found to be superfluous when observing the characteristics of communication [5]. However, the prosody (or melody) within language becomes increasingly more necessary when observed as poetry or music (Perlovsky, 2013:3), and one of the key aims of the present paper is to address the shared characteristics exhibited by the means of communication that exist between primitive language and music, in what is herein referred to as an emotional-linguistic spectrum, or ELS (figure 6).

---

[5] This statement is predicated on a quantitative analysis of emotional linguistic research against that which pertains to social, evolutionary or ontological fields.

If we begin by considering pidgin language, or what we will deem to be at the opposite end of the ELS to that of music, we find language employed for the sole purpose of exchanging information. Pidgin languages are purposely simplified for communication between groups who do not share a common language. In some cases they will be based on an pre-existing language such as English or Chinese, but again, will consist of a reduced subset of the vocabulary. In either scenario, a lack of morphophonemic variation exists, and a concern for intonation or articulation would seem impractical.

Being the basis for pidgin, native languages subsume any such emotionality, and thus occupy the adjacent point in the ELS. However, as the prosody of native languages remains subjective, not all can occupy the same definitive point on the scale. For example, Finnish speakers find their native tongue to be comparatively unemotional by comparison to English (Erkkilä, 2014); the premise of which based on the lack of aspiration innate in Finnish plosives, which in English is frequently emphasised.

Auxiliary languages, or *auxlangs*, would arguably be placed alongside native languages as they are also subjective in their emotional content. Auxlangs, such as Zamenhof's Esperanto or Schleyer's Volapük, are similar in their ideology to that of pidgin languages, in presenting a means by which to communicate where a common native language does not exist. However, whereas pidgin develops naturally to resolve a need for a common language [6], auxlangs are created artificially with the sole purpose of facilitating international communication, thus requiring no such impetus. Whilst it can be argued that emotional differences could be noticed between auxlangs, it is not the intention of this paper to scrutinise such languages on a case-by-case basis.

Unlike the aforementioned examples, artistic languages, or *artlangs*, are created with the purpose of imparting aesthetic pleasure through vocabulary, and much like Tolkien's Elvish and Orwell's Newspeak, often accompany a

---

[6] For example, Hawaiian Pidgin originated as a result of non-English speaking immigrants settling on the islands who required a common language whilst working in the plantations alongside the English speaking natives.

fictional world made manifest by a novel, film or TV series. In the same way that one observes a foreign language, artlangs are perceived in ignorance of their lexical context; the result of which denoting a theory that suggests the auditors concentration being further biased towards the *sound* of the language (figures 4a & 4b). It is therefore an objective of this paper to test such a hypothesis in showing the observed emotionality of artlangs to be comparably greater to that of native languages.
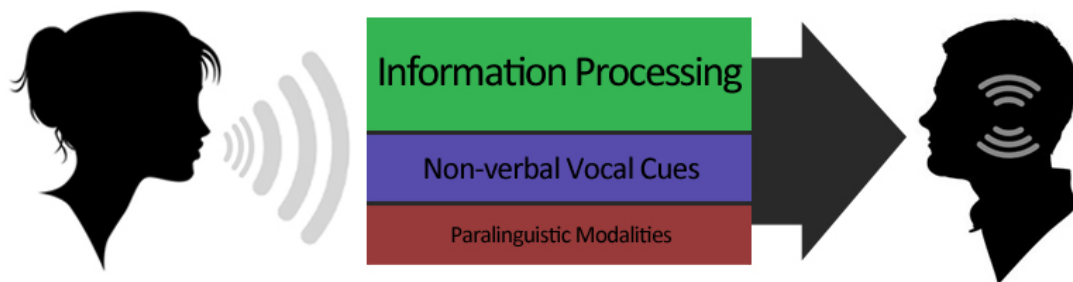

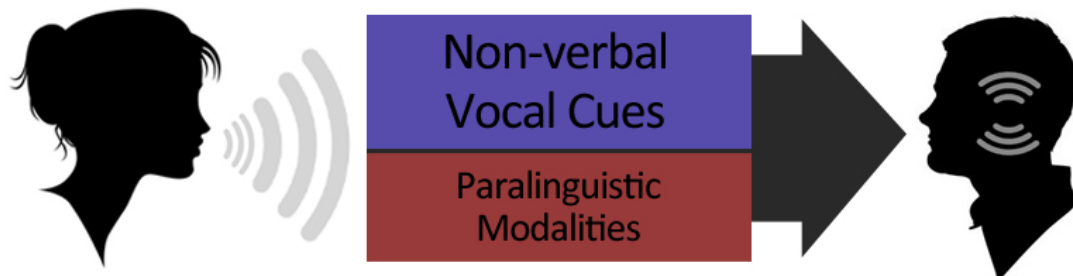Figure 4a - Speech perception of native languages

Figure 4b - Speech perception of artistic (or foreign) languages

Whilst the artlang creator has a unique freedom over rules pertaining to native language, the same can be said of the poet. By arranging literature within the context of prosody (sound and rhythm moreover the meaning), the poet can write freely without giving respect to social or historical contexts to which the words belong (Leech, 1991:5). Furthermore, the concept of artistic language appears to have long since been a consideration of poets, as Thomas

Gray once wrote, "everyone that has written [poetry], has added something by enriching it with foreign idioms and derivatives; sometimes words of their own composition or invention" (Leech, 1991:15). A good example of Gray's observation is Lewis Carroll's *Jabberwocky*, where we find a number of "nonsense" words which, despite their lack of semantic context seemed "very pretty" and "filled [his] head with ideas" (Carroll, 2010, 64).

'Twas **brillig**, and the **slithy toves**
Did **gyre** and **gimble** in the **wabe**;
All **mimsy** were the **borogoves**,
And the **mome raths outgrabe**.

Beware the **Jabberwock**, my son!
The jaws that bite, the claws that catch!
Beware the **Jubjub** bird, and shun
The **frumious Bandersnatch**!

Figure 5 - First two stanzas of Carroll's Jabberwocky (nonsense words in bold)

Whilst Carroll's experiment in poetic language may today be seen as an attempt at artistic language, it was written at a developmental stage of *sound poetry*, during which the non-semantic and acoustic properties of language were isolated and explored for their own sake (McCaffery, 2008). By limiting syllabic content and focusing on pitch and amplitude variation, sound poetry reduced meaning and structure in an attempt to unite literature and music. This movement gave rise to such works as Zaum [7] and Schwitter's Primal Sonata - or what was referred to as "the artistic moulding of nonsense" (Scobie 1974: 222). Despite being a short lived example of experimental linguistics, similar characteristics are also found in the primitivist chants of indigenous

---

[7] Zaum is the collective name given to the linguistic experiments (also referred to as Zaum poetry) in sound symbolism carried out by various Russian Futurists, the most notable of which being Khlebnikov and Kruchenykh. Their work was created without meaning to allow for fuller expression, which the Futurists considered to be a bind in native languages.

settlements, (non-referential) performances of avant-garde theatre, and the chants of spiritual or religious practises such as glossolalia (Ovadija, 2013), or what is more frequently referred to as speaking in tongues [8].

The scale then concludes with the melodic vocals associated with song, or musical vocalisation. Traditionally the lyrics of a song will share the syntactical rules of its respective native language, but also utilise two additional acoustic cues; rhythm and melody. Studies have shown emotional arousal to increase further during non-verbal vocals (Loui et al., 2013), suggesting that the timbre (or tone) of musical vocalisation is more significant than the lyrical aspects in satisfying emotion. Irrespective of the emotional trigger, decoding accuracy for emotion communicated through music has been shown to be above chance performance levels (Ilie, 2006:320), thus placing it at a point on the scale above the aforementioned modes of communication.

The following diagram summarises the aforementioned elements of the emotional linguistic spectrum, indicating the hypothetical levels of perceived emotion. Although other areas will be discussed, the key focus of this research will isolate the native, non-native and artistic languages.
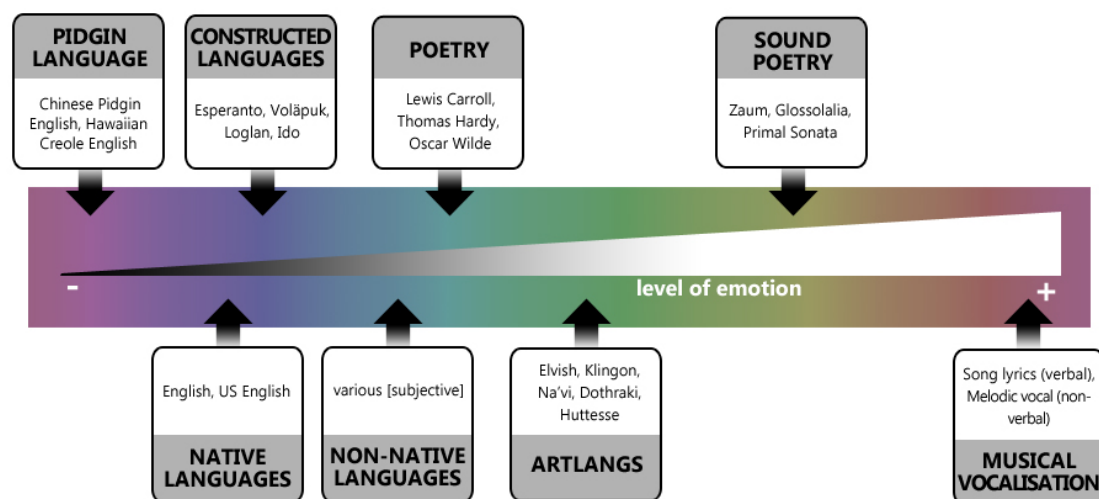

Figure 6 - The Emotional Linguistic Spectrum (ELS)

---

[8] Glossolalia is the term given to verbal behaviour consisting of the phonemes consistent with the speakers native language, which is reconstructed, somewhat psedogrammatically, to resemble structures similar to sentences.

Evaluating the Perceptual Linguistic Emotions

Emotions in the brain are used to communicate instinctual needs, and the brain evaluates emotional signals for the purpose of satisfying these needs. Lower level or "bottom-up" responses driven by neural primitives are thought to be employed for such purposes, which do not involve the higher cognitive (or "top-down") mechanisms associated with concept or rule-based systems, and as required by pidgin or native language communication (Perlovsky 2010:10).
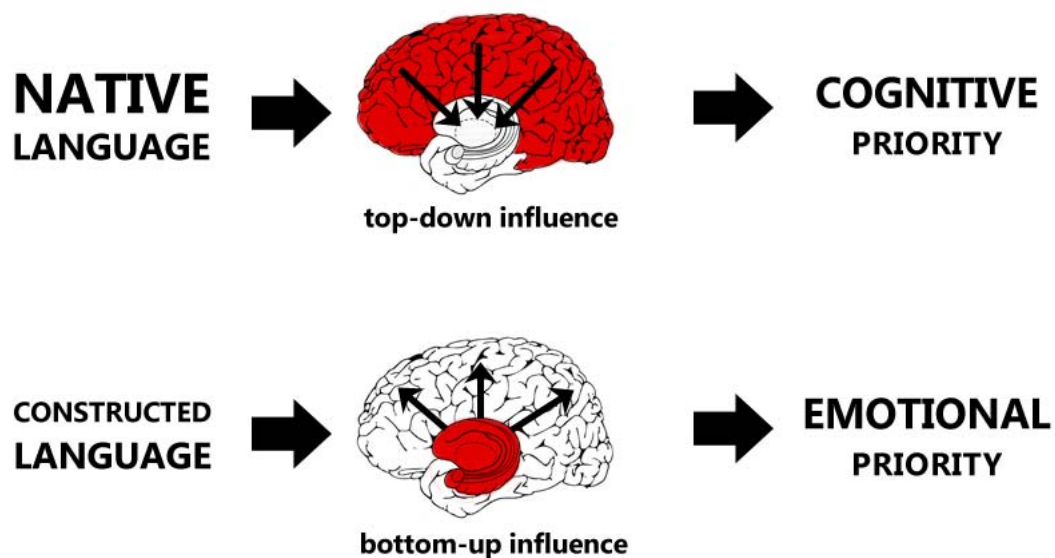

Figure 7 - The response mechanisms of language in the brain

As a result of these responses, the further we move along the ELS, the more concentrated the language becomes in receiving preferential attention from the areas of the brain responsible for the evaluation of emotional responses. All of the questions posed by this paper are therefore predicated on these observations of the lower and higher level response mechanisms, and the conjectured bias towards one or the other.

In the first instance, we return to Saussure's model of concepts and sound-images. The recognition of the concept is universal, whilst the sound-image

becomes a subset of the concept defined on a language level. Both will occupy cognitive mechanisms in the brain. That which then follows these mechanisms is then required to quantify and appropriate the subsequent emotional response. The areas responsible for doing so, occupy the lower level part of the brain, thus becoming subject to interpretation. That is to say, memories that involve aspects of emotion are a combined function of the amygdala and the hippocampus, both of which form areas of the same lower level, or the aforementioned limbic system. Therefore, whilst emotional responses to any given sound-image may be shared between individuals, the means by which the response is perceived remains subjective, and therefore cannot be attained demographically. The diagram below illustrates the proposed extension to Saussure's model.

EMOTIONAL PERCEPTION personal	HAPPY (cute, fluffy) ANGRY (pest, vermin)	UNHAPPY (winter, cold) RELAXED (artistic, beautiful)	PEACEFUL (skylines, holiday) AGITATED (heights, fear)
SOUND IMAGE language	"rabbit"	"bare tree"	"skyscraper"
CONCEPT universal			

Figure 8 - The proposed extension of Saussure's concept
and sound image model

Whilst the cognitive areas of the brain used for processing language continue to develop into adulthood, the amygdala is considered to be largely developed at birth. However, despite its early development, the amygdala maintains neurological paths to the cortex, which in turn enable the exchange of

emotional signals. The resultant responses then cause an affective state that manifests itself in a number of physiological changes; one being the voice production process, from which the auditor can then infer the respective emotional state. For example, the arousal associated with an elevated mood would cause an increase in the perceived vocal pitch, along with a decreased syllabic duration to match the heightened respiration. Therefore, whilst little is currently known of the cognitive interaction between language and emotion, we are aware that it has the potential to encode emotion.

However, in view of the limited significance that emotional encoding is given in the exchange of conceptual information, the influence is frequently overlooked. To such an end, we begin by analysing the emotional potential of native language. Whilst the amount of emotional feedback may be minimal, that which is deduced from measuring the responses may then be used to infer the results from subsequent bilingual participants, and potentially the same of non-native and artistic languages.

Experiment 1: Monolingual Emotion

The means by which information can be collected for such a study can easily invoke further subjective interpretation. The stimuli used to trigger affective responses are likely to have different results based on the individuals emotional recollection of the concept. For example, the concept of a "dog" is likely to have a positive or neutral response for most, whilst others may have a negative reaction owing to a pathological fear. So despite there being no intrinsic relationship between the sound-image and its relative concept [9], the response becomes, in part, subject to an affective connotation. However continuity does appears to pertain to the stimulating medium. Regardless of whether a participant views an image, hears *or* reads the word, a majority of

---

[9] As Saussure states, there is nothing inherently doggish about the word dog, and nothing but the lexical convention that keeps us from calling them "blogs" (Wiese, 2004:124).

the same neurons fire in the brain, thus the same emotional response is likely to result (Koch, 2005, Heiser et al. 2003, Lauer 2007, 149).

Despite the innate arbitrariness, Whissell, among others, saw an opportunity to build upon this existing framework by conducting surveys that would assign emotional "scores" to individual words and phrases. The result of this work was a *Dictionary of Affect in Language* (or DAL), consisting of approximately 9000 entries [10]. Each word is rated using Osgood's two-dimensional model of evaluation and arousal, or what Whissell refers to as pleasantness and activation (Whissell, 2014). On a scale between 1-20, each value represents the polarity and concentration of the perceived emotion; the mid-point of which representing a neutral response and an average activation. The word "tree" for example, is considered to be moderately pleasant (15), with a average activation (10). Whissell's interpretation of such a score states that the word tree *sounds* particularly pleasant, but with neither a strong or weak emotional impact. The word "hate" however, shows a comparatively unpleasant score (7), coupled with a stronger emotional activation (15). The table below shows example results from Whissell's research.

	Pleasantness	Activation		Pleasantness	Activation
love	16	16	hate	7	15
sleep	15	8	run	12	17
hugs	15	10	fight	7	15
up	11	13	down	7	8
happy	16	17	sad	7	9

Figure 9 - Example words from Whissell's Dictionary of Affect on Language

---

[10] First published in 1986, Whissell's DAL is a project of continued research that results in the frequent addition of entries. The figure shown is considered approximate to the time of writing.

From the examples it could be argued that the denotative meaning remains a consistent influence; "run" is suitably active in respect of the associated concept, whilst "love" and "hate" are opposed in pleasantness, as are "happy" and "sad". Whissell then examines the relationship between each score and the occurrence of individual phonemes, and using the Pearson correlation coefficient (*r*), categorised the graphemes and digraphs into their respective divisions, resulting in the emotional compass shown in figure 3. The statistics were later cross-validated by repeating similar surveys using what Whissell refers to as nonsense words, with results shown to be correct at better than chance levels (Whissell, 2014:12).

Further research conducted by Auracher, Albers, Zhai, Gareeva and Stavnijchuk, tested the cross-cultural validity of Whissell's claims by observing similar characteristics exhibited by a number of different language [11]. However, the results contradicted Whissell's results in a number of areas. For example, Whissell's results show English participants rating plosives as nasty or unpleasant (Whissell, 2014:24), whereas Auracher et al showed them as pleasant (Auracher et al., 2010:21) [12], thus indicating that encoded emotionality remains language specific. Taylor and Taylor also conducted a study consisting only of nonsense words, formed using sounds common to the participants native language. Whilst consistency was also evident in this case, it too remained language specific (Taylor & Taylor, 1965:413).

So whilst these results show affective consistency in the context of the native speaker, denotative influence also appears equally consistent. So whilst lower level (emotional) mechanisms can influence the information processing of the higher level responses, conjecture remains when establishing the concentration of this influence. Support of this theory is evidenced in Jespersen's case studies, through which he suggests the familiarity and habit

---

[11] Auracher et al. selected poems from each native language (German, Russian, Chinese and Ukrainian) of the participants, all of which were analysed to calculate the ratio of plosives to nasals. The highest and lowest of each selection were then used to determine the above results.

[12] The results of the experiments undertaken by Auracher at al. showed a "universal tendency" to use plosive phonemes in expressing happy and active feelings.

of native language attunes our cognitive processes to instinctively accept words as adequate to express the ideas they stand for (Falk, 1992). So at what point does this phonetic habit preside over the influence of lower level responses? Furthermore, given that constructed (or nonsense) words do not occupy the processes required by native languages, it is logical to assume that the lower level areas get preferential influence over any subsequent affective response? For example, when reading constructed vocabulary, we are aware that any lexical analysis would be futile, thus the phonemic structure of the vocabulary remains the primary influence from which an affective response can result [13]. Therefore, if constructed languages are more susceptible to emotional influence, surely nonsense words would be the ideal means by which to determine the most accurate response?

Furthermore, when considering the level of emotional engagement in native language, rather than making use of words that have limited synonymy, it would seem logical to focus on those whose definition has more semantic variation. That is to say, assuming one *can* disassociate the denotative influence, when asked to choose a word that best describes a specific emotion, will the user will find themselves drawn to specific phonological sequences? Does the word "beautiful" reflect the aesthetics of the implied emotion? Or is "gorgeous" or "lovely" more fitting? If a pattern can be established from the responses, it is plausible that emotion can be decoded from specific phonemes within the lexicon of native languages. Conversely, should a pattern not be evident, it is plausible that our higher level responses have presupposed a suitable response, thus operating on the basis of pseudo-emotional stimuli.

However, if words were *created* to represent emotions, thereby diminishing the denotative influence, would users find themselves drawn to words containing specific phonemic attributes? Would the consistency show a more markedly difference to that of native language results?

---

[13] Whilst lexical consideration would appear not be a factor when reading constructed words, the phonology used may resemble words contained within the participants native lexicon, and thus may remain an influence.

In order to further understand the emotional influence of a native language compared to that of a constructed language, surveys were written to query the response to the same emotion in both cases. Statements were prepared, each with a missing word, denoting active, neutral and passive aspects of both positive and negative emotions [14]. Users were then asked to choose a word which they felt most appropriate for the context of the sentence, whilst making every effort to ignore the respective semantic connotation. Each of the statements gave the option of three words; one with a high concentration of plosive phonemes (disgusting, outstanding), another biased towards nasals (animosity, phenomenal), and a third with a balance of vowels and consonants (offensive, elegant). The purpose for using such groups was to draw attention to the phonemic content by means of pronunciation. By suitably diversifying the phonemic features of the words, the easier it is to divert the users attention to the sound moreover the definition, thus promising the most accurate results. Furthermore, by targeting English monolinguals, participants would have no prior knowledge of foreign languages, and thus could not influence the results.

In this first group of questions, users were asked to choose words to complete sentences of both positive and negative emotional context. The results (figure 9) show an even dispersion over the three aforementioned phonemic groups, suggesting that despite efforts made to ignore semantic associations, phonemes have little, if any, influence over the choice of vocabulary in any given context. Therefore, Jespersen's hypothesis of habit appears plausible at this stage, and that familiarity does indeed take precedence over any potential emotional encoding.

---

[14] Active emotions are those which intrinsically cause motivation or vigour such as excitement or anger, whilst passive emotions incorporate inactivity or idleness, such as boredom or relaxation. Neutral emotions are those which suggest neither inactivity or motivation, such as beauty, or pleasantness.
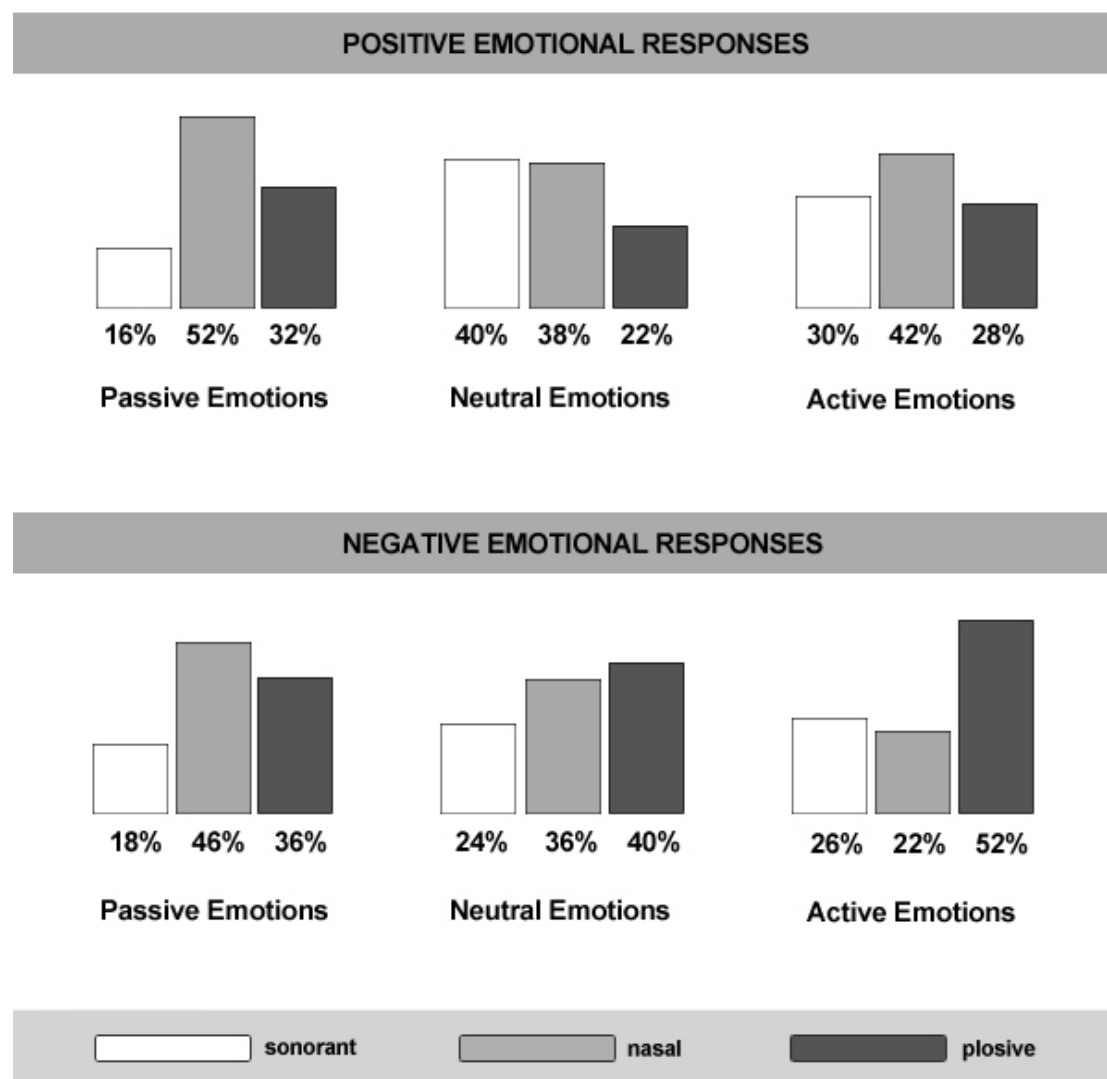
Figure 10 - Results of English words chosen by English monolinguals to best represent positive and negative emotional context

The two areas of deviation from an otherwise uniform set of results are shown in the preference of nasal syllables to represent passive emotions in the positive group, and the plosives in the negative active group. Whilst both could be attributed to semantic familiarity, the latter has been noted in previous emotive speech studies. Fónagy postulates that negative speech patterns are considered biological metaphors of nausea and coughing, and that plosive phonemes represent the rejection of harmful substances, characterised by the same increased muscular tension (Fónagy, 1997:7). Conversely, by expressing positive emotions, speech organs become relaxed,

thereby allowing sonorant syllables to become more enunciated (Fónagy 1991:495). Whilst this may be true for the parameters of pronunciation in both respects, the above results only concur with the first statement.

In the second part of the survey, further multiple choice questions were given using the same phonetic groups to best describe a given emotion. However, the same methods as those employed by Sapir and Whorf were used, whereby the native English words were replaced with constructed examples to further reduce the influence of the respective concept. In the interest of ensuring that the constructed words could be read accurately, English phonology was adhered to, but in such a way as to avoid phonemic similarity with native words of a similar affective definition. For example, use of the stem "ther-" could invoke the word "*ther*apeutic", in the same way "dis-" may suggest "*dis*turb" or "*dis*gust".

From the results (figure 11), a more defined distribution pattern is immediately noticeable. The positive results collectively show a mean deviation increase of +25.85 from the first test (*M*=33.3, *SD*=10.9), to that of the second (*SD*=24.9), and an increase of +21.16 in the negative results from the first (*SD*=11.57) to the second test (*SD*=31.68). A clear division between the passive responses, and those from the active and neutral is also noticed similarly across both positive and negative responses. Plosives are shown to best reproduce both active and neutral emotion, whereas sonorous words are noted as preferable for passive responses. Furthermore, despite showing a similar distribution pattern, a greater preference is shown for sonorous words in the *positive* passive group over the respective negative group. In the same way, the plosive response for active emotions is shown to increase in the *negative* group.

Collectively these results show that a negative context attracts a greater *active* response, whilst a positive context attracts a greater *passive* response, thus denoting that words such as "hate", "disgusting", "therapeutic" and "eternity" amplify the emotional response in a native context. The responses at this stage

also further support Fónagy's aforementioned hypothesis of muscular tension characteristics in phonology.

POSITIVE EMOTIONAL RESPONSES

72%   24%   4%
Passive Emotions

20%   16%   64%
Neutral Emotions

16%   24%   60%
Active Emotions

NEGATIVE EMOTIONAL RESPONSES

56%   40%   4%
Passive Emotions

18%   16%   66%
Neutral Emotions

8%   4%   88%
Active Emotions

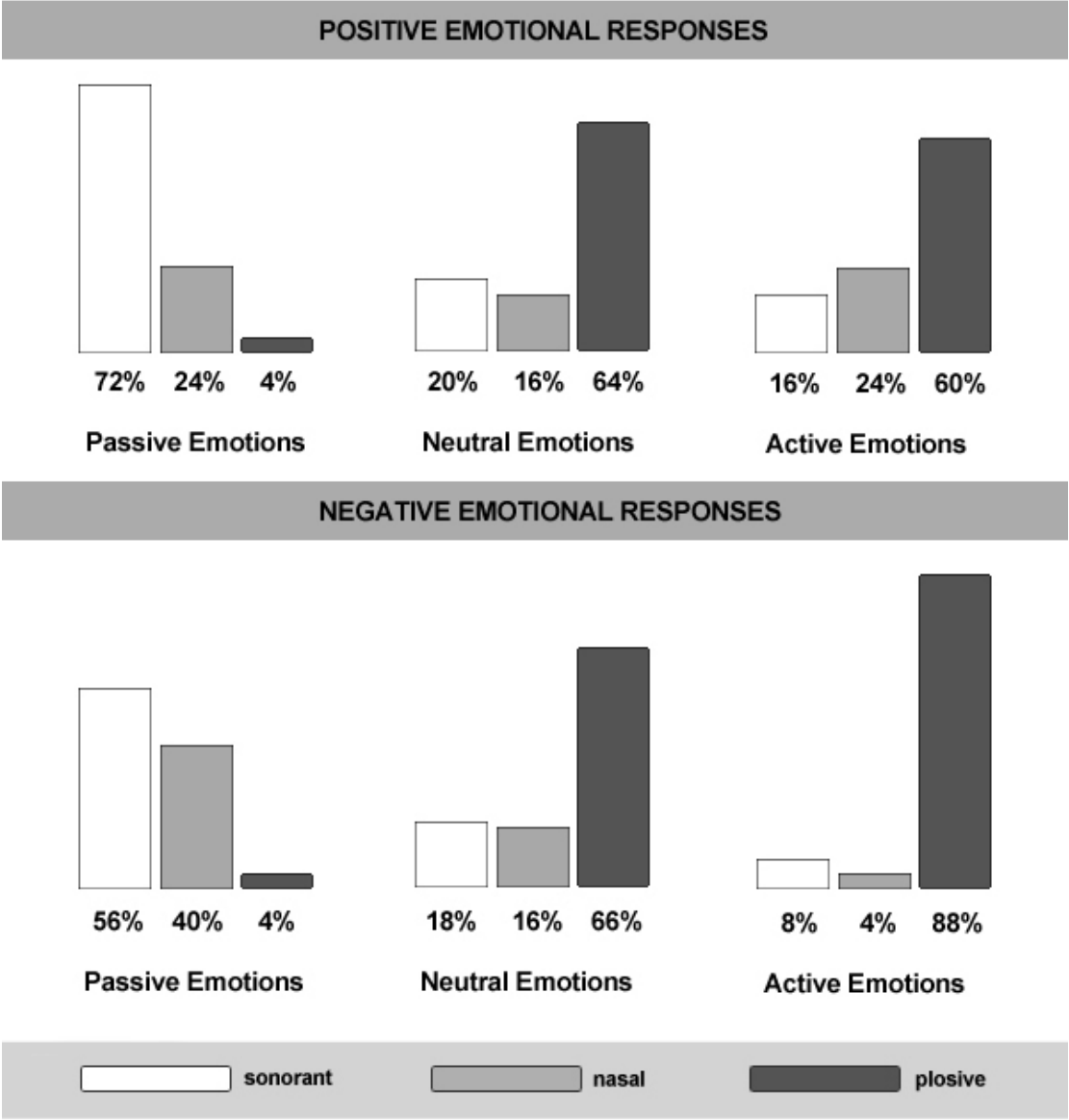sonorant   nasal   plosive

Figure 11 - Results of reading constructed language words of both positive and negative emotional context

Regardless of whether a body of text is of native or constructed origin, the same area of the brain, known as the angular gyrus, will be engaged for its comprehension. Therefore it seemed prudent to pursue areas of testing that would engage additional parts of the brain, thus giving additional perspective

to the previous results. Whilst little is known of the creative processes that occupy the brain, the cognitive function required to *create* constructed, or nonsense words, would employ areas of the brain in addition to those required for reading. To this end, a further test was appended to the above, in which participants were asked to create their own words for an emotion of their choosing.

These questions, as with the previous, were interspersed in the study so as to reduce the potential influence from previous constructed and native examples. Responses were categorised into the required groups of sonorant, nasal and plosive based on the frequency and position of the respective vowels and consonants. For example, "ferath" and "pelarin" were categorised as sonorant words, given that both have a CVC structure and use consonants from different points of articulation. Conversely, "bogartin" and "peklufig" were considered plosive responses due to the consonant grouping and frequency of plosive phonemes. However, where plosives occurred adjacent to a vowel, the articulatory position of the vowel was considered when assigning a category. For example, a front vowel in English phonology (i, e, y) will typically reduce the aspiration of a plosive, whereas a back vowel (u, ɑ, o) will amplify the aspiration. Therefore, words such as "pulikaraad" (p**u**likarɑd), whilst containing a CVC structure, increased the effects of the plosive aspiration, and thus were categorised as a plosive response.

As anticipated by the previous results, the passive and active groups in both the positive and negative sets saw a further increase of sonorant and plosive responses respectively, with the deviation of the conjectured progression shown in the positive neutral responses. Rather than maintaining the plosive responses of the second test, the results indicated a sonorant syllabic preference. The respective negative group maintained an increase of plosives, thus further separating the two sets. The results of the survey are shown in the following diagram (figure 12).

POSITIVE EMOTIONAL RESPONSES

100% 0% 0%	90% 2% 8%	4% 0% 96%
Passive Emotions	Neutral Emotions	Active Emotions

NEGATIVE EMOTIONAL RESPONSES

72% 28% 0%	4% 6% 90%	8% 0% 92%
Passive Emotions	Neutral Emotions	Active Emotions

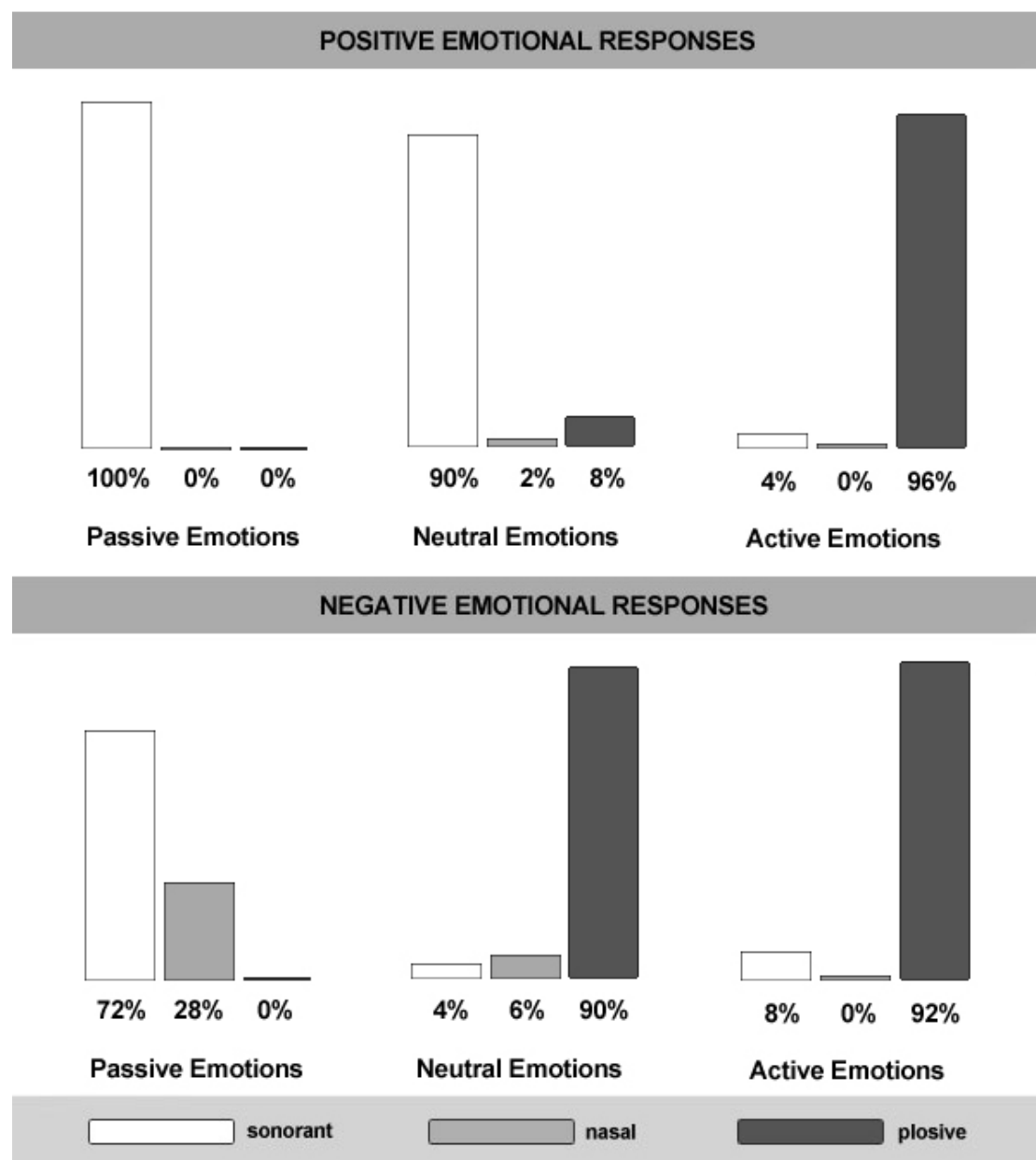sonorant   nasal   plosive

Figure 12 - Results of user written constructed language words
of both positive and negative emotional context

Collectively the tests show a continuous increase of consistency that dictates
sonorant words are best used to represent passive emotions, and plosive
words to represent active emotions (figure 13). With the exception of the
positive active set, all results show an exponential change in the plosive
responses. This suggests that plosive phonemes are easily recognised in their
role of active emotional encoding. Other responses in the passive and active

sets give the same indication, although on a more uniform scale. Whilst similar trends are shown for the constructed writing responses in both positive and negative neutral sets, the overall results do not appear as conclusive by comparison to the above. This therefore suggests that the bipolar aspects of valence aids the user in identifying the phonetic emotional stimuli.
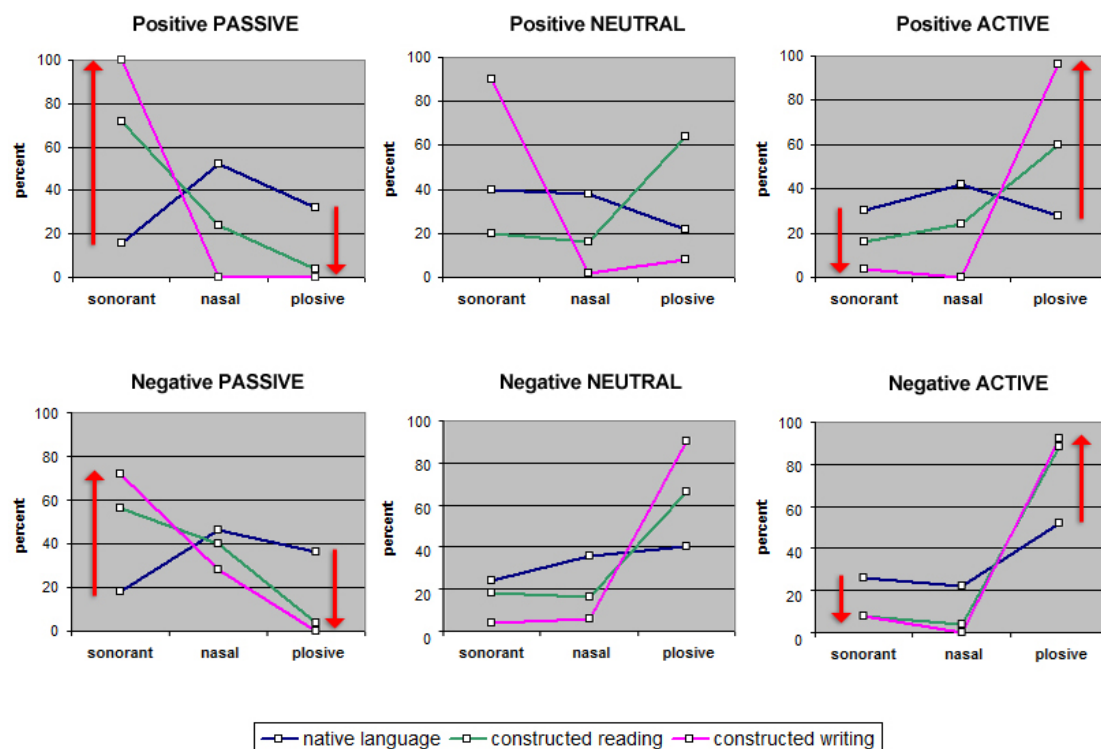

Figure 13 - Stacked results of both positive and negative
monolingual emotional test responses

These results also further evidence those obtained by Auracher et al., having also concluded there to be a tendency towards plosive syllables in the expressing of active emotions (Auracher et al. 2010:21). The same tests also show a preference for nasal sounds to induce negative and passive emotion. Whilst a decrease in this respect is noted in the above responses, they remain the results that show the least amount of diversity over the three tests. Finally, despite the consistency shown in the results, Auracher et al. made no explicit

references to, nor made exclusive use of sonorant syllables, thus no comparative analysis cannot be drawn.

Whissell's results also show nasal phonemes to be regarded as passive, although the alveolar nasal (n) is shown to be associated with *negative* passive connotations. However, the plosive test results from the present study do not corroborate with those shown by Whissell. As indicated in the aforementioned emotional compass (figure 3), all plosives are characterised as "nasty", along with a designation for both active *and* passive associations. Whilst both results agree that the voiceless alveolar plosive (t) and voiced velar plosive (g) are predominantly active phonemes, there is no association shown for positive active emotion, as strongly indicated in the present study. The voiceless velar plosive (k), although also categorised as nasty and unpleasant, is shown to be a passive consonant, again contradicting the uniform decrease of plosive in both passive sets shown above.

Crystal conducted a similar analysis using words that scored highly on a national survey that asked participants to vote for what they considered to be the most aesthetically pleasing English word. From these results it was concluded that one could create phonaesthetically pleasing new words by combining, among other attributes, an equal combination of consonant and vowel articulation, whilst avoiding the use of consonant clusters (Crystal, 1995:11) [15]. Collectively, the above results agree with Crystal's findings, assuming the end result would be classified as an aesthetically pleasing *passive* word. However, a majority of the words used for the analysis, whilst denoting positive valence, did not represent specific emotions, or a level of arousal (e.g. lullaby, melody, chalice, amaryllis etc.).

Despite the results gleaned from the reduction of lexical influence, the question remains as to whether the users native lexicon remains as a basis from which the affective response is perceived. That is to say, do native English speakers find words such as "lovely" and "therapy" positively phonetic

---

[15] A summation of Crystal's analysis is found in the concluding words of his paper where he states that "a romantic poem about London Underground stations [is likely to include] *Pimlico* and *Colindale*, but exclude *Goodge Street* and *Wapping*" (Crystal, 1995:12).

purely as a result of Jespersen's theory of habit, or is there an underlying universal phonetic aesthetic? In the case of the former, how does this affect participants familiar with more than one language? And what influences are responsible for phonetic bias towards the lexicon of one language over another?

Experiment 2-1: Bilingualism - an aesthetic extension?

Whilst the use of constructed words may reduce the association of semantic reference, the same phonology is represented. As a result, this may limit the aesthetic potential of phonemes used in other languages. For example, Finnish uses 30% more vowel and diphthong phonemes by comparison to English, some of which may be found aesthetically pleasing by non-native speakers. An example of which being JRR Tolkien, who made extensive use of Finnish phonology and the associated vowel harmony in the Elvish languages used in his fictional works of *The Hobbit* and *The Lord of the Rings*.

However, this is not to say that a larger inventory of phonemes will necessarily denote a more definitive emotional lexicon; this would be better defined by the diversity of the languages phonology. Whilst other languages share a number of phonemes with English, the frequency and conditions by which consonants are combined differ, thus resulting in the aforementioned paradigm of aesthetic interest. Portuguese for example uses the same /l/ and /n/ phonemes found in English, but are further subject to palatisation, under which conditions become /lh/ and /nh/. These digraphs are pronounced as palatal lateral approximants (ʎ), and the palatal nasals (ɲ) respectively; both of which are unused in the English language.

A number of languages also change the pronunciation characteristics of phonemes. Although represented by the same grapheme, plosives used in Finnish undergo a process of gradation, and are distinguished by "strong" and "weak" grades. Where a single grapheme is used to represent a plosive

phoneme, the weaker grade applies, for example *papit* (tr. priest) and denotes a reduction of aspiration, or voicing. Under conditions where more aspiration is required, the same grapheme is doubled, for example *pappi* (tr. priests), thus becoming the stronger grade. Whilst this process isn't established in the same way, a similar process is evident in the English language, for example when pronouncing "**p**air" and "s**p**are". However, in the case of Finnish consonants, it is further possible for the aspiration to be removed altogether; an example of which can heard in the voiced dental plosive (d), which when weakened becomes phonetically characteristic of its voiceless counterpart (t) [16] (Suomi et al. 2006:159). Whilst the same can be found in English words, such as "a**p**t" and "do**c**tor", the aspiration becomes assimilated solely as a result of the neighbouring plosive. By comparison, Finnish has a far more complex morphology that governs the gradation of consonants, which in turn pervade the language.

Despite the comparative similarities between English and foreign languages, unrecognised phonology can readily invoke aesthetic interest, thus becoming a factor for aesthetic consideration. As previously mentioned, native Finnish speakers have been known to prefer the English phonology by comparison to their own, due to the increased frequency of aspirated plosives, which permits more syllabic emphasis that in turn increases the affect in everyday speech (Erkkilä, 2014). Furthermore, given Tolkien's affinity for Finnish, and the subsequent use in his own artistic languages, this same phenomenon may also explain why native Finnish speakers do not experience the same aesthetic pleasure from the Elvish languages as English speakers.

Therefore, in order to alleviate the possibility of biased aesthetic interest by using the same monolingual participants as used for the first experiment, bilingual speakers were targeted for the next stage of the research. Those that responded were asked that they use both languages daily, and where possible,

---

[16] Whilst this is considered to be the typical Finnish pronunciation, dialectal realisation causes this to vary.

for both to be used in similar circumstances [17]. The logic for using such participants was based on the well-reasoned assumption that being familiar with, and subject to, a more diverse morphology, the user would have experience of a wider phonemic arsenal, of which could potentially invoke emotional responses. Furthermore, the same user could objectively compare one language with the other, thus giving greater scope for their respective aesthetic strengths and weaknesses.

The initial part of the survey established the participants native tongue, and that which they considered to be their second, or auxiliary language. A majority reported to have learnt both languages from early childhood as a result of being subject to bilingual parenting; others were learnt as a requirement of relocation, again from family or school, and subsequently improved to a level of native fluency. Of the responses received, seven individual languages were represented for both native and second languages - English, Spanish, Greek, Dutch, French Russian and Portuguese. The function of the second language was used mainly for family communication, and for those who learnt their second language whilst in education, also reported to be using the same for work, everyday use and social activities (figure 14). Participants were also asked whether they would choose to speak one language in preference to the other, and if so, the circumstances for which this would occur. Some participants stated that a specific language would be spoken for the purposes of formality, whilst others stated that preference would be given as a means for privacy in an environment where discretion was a desirable factor. However, the most frequent response was that pertaining to improved emotional communication. Whilst the reason for such a decision may be related to the language lexicon moreover a phonemic context, the remainder of the survey would focus on methods of retrieving and analysing the influences behind this response.

---

[17] A majority of those who took part in this study reported to use their languages for family, friends and in their working environment. Whilst most stated that their native tongue was used for communication with their family, thus incorporating emotional context, the same users also reported that their second language was used when conversing with friends or partners, which would also involve the same level of emotional intimacy.
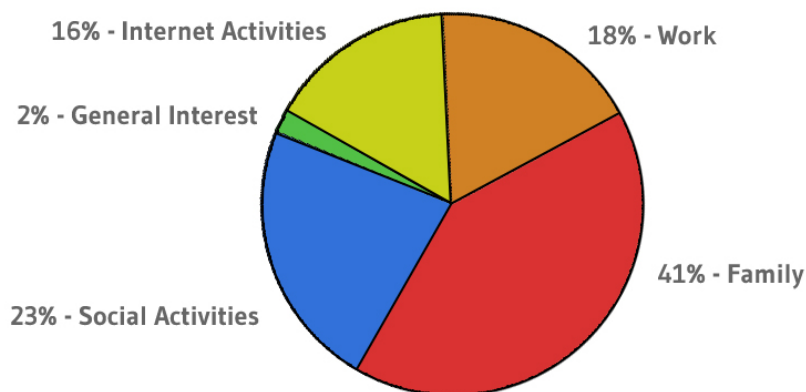
Figure 14 - The distribution for the use of (bilingual) second languages

The following part of the survey asked participants to rate their languages in terms of emotion and aesthetic. The purpose of doing so was to initially identify how bilinguals rated one language against the other, and secondly to compare those results with speakers of the other languages represented in the experiment. The results shown below (figure 15) indicate the mean response to the given affective criteria, and show the combined results from both native and non-native speakers of each respective language.

Greek is shown to be both the most expressive and emotional, and based on information given by (native) Greek speakers, this is due to there being, or what the participants believe to be, a greater amount of words available to express concepts that would otherwise not be possible in English. However, whilst it is accepted that the Greek lexicon includes words that exemplify concepts found in English, that is not to say that English itself is limited in communicating Greek semantics. For example, the words *agápe* and *éros* both extend the meaning of the word "love", although the same concept can be expressed in English by use of the appropriate adverb; in this case "charitable" and "intimate" [18]. So whilst not all Greek words can be quantitatively

---

[18] *Agápe* was first appropriated in Christian theology to describe the love between Christ and humankind, and the reciprocal brotherly love for God. In modern Greek, this has been extended for use to describe the love in a charitable context when addressing a person or group. So whilst the same word can be used to describe a number of alternate concepts, appropriate vocabulary exists in the English language to effectively describe any given concept. *Éros* conforms to the same paradigm.

translated in a single word, it is nonetheless possible to convey the same concept cross-culturally.
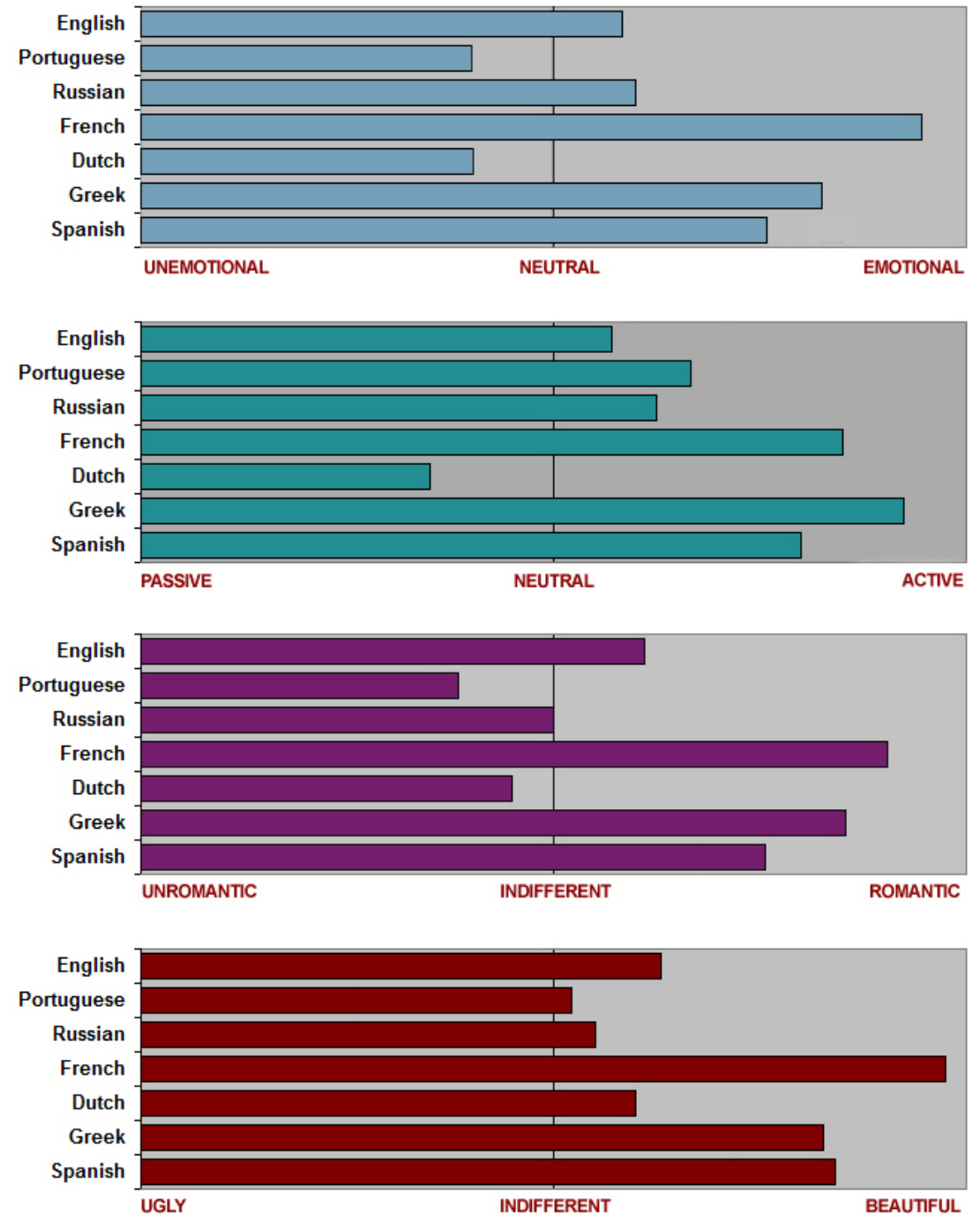

Figure 15 - Results (averaged) of emotional scoring by bilingual participants to given affective criteria

The French language was shown overall to be the most beautiful and romantic language, and from interviews with two French speaking bilinguals (one native English, the other native French), similar insights were found regarding the reasons for the above results. In the same way that Germany is associated with hearty meat-based cuisine and beer, so to is France associated with designer fashion, perfumes and fine wines. However, despite the stereotypical associations, the cliché remains an influence for many aspects of the French culture - including the language (Lasbareilles C-L, 2014). As such, the spoken word appears to induce similar emotive states to that of the Eiffel Tower, or the concept of a romantic break, and that native and fluent speakers become something akin to an advocate for their respective culture.

In addition, comments were made regarding the fluidity of the French language when compared to Italian and English, and in relation to the phonemic structure, or what is collectively perceived as "choppy" (Lasbareilles C, 2014) [19]. Analysing native French speech patterns, we find the variation of the fundamental frequency (F_0), or *jitter*, vastly reduced by contrast to both Italian and English. Can we therefore assume that greater emotional emphasis is attributed to other phonemic characteristics? Can this diversion of emphasis somehow explain the results shown above?

The perceived pleasantness of the French language appeared to be greatly supported by the disliking of the aforementioned intonation qualities, but also in addition to an accompanying innate "falseness", that makes languages that employ such characteristics more difficult and uncomfortable to hear (Lasbareilles C-L, 2014). The metaphor given in this instance was that of an airline stewardess or front-desk receptionist whose foremost responsibility it is to instil positivity during customer interaction. Done so by employing an elevated F_0 and increase in jitter, the dislike occurs not only from the undulating intonation, but from the behaviour of naively concealing the facade beneath.

---

[19] One such metaphoric comparison between English and the French language was that of snoring and purring (French being likened to the latter). Whilst they are similar in their biological means of emitting noise (as is the English and French language in that they both use a phonemic structure), one is nonetheless a nicer sound.

A comparable example established in emotion, would be that of greeting a friend from whom one has been separated for a time, suitably long enough to greet with an embrace, which is often met with the same elevated pitch and increased intonation. This however, is not the case with the French language. Rather than an overt display of arm-waving and high-pitched greetings (or what was referred to as a state of being "linguistically eccentric"), the F_0 is uniform regardless of the perceived emotion or its respective valence (Lasbareilles, C-L, 2014). In opposition, native French speakers will elongate and soften vowels and digraphs, again with minimal inflection. Furthermore, speakers will lower the overall voice register to compensate the level of emotion appropriate for the situation, or to sum, the lower the F_0, the more emotion is implied.

When the interviewees were questioned regarding the perceived emotionality for the continued use of both the English and French language, it was stated that adapting to pitch change is done so largely out of cultural acceptance, and that the perception of emotion is often overlooked. Whilst similar circumstances are likely to be found in other dialects and languages, suffice to say that a majority of the emotion encoded in similar acoustic cues are either lost or misinterpreted.

Experiment 2-2: Analysing the Collective Aesthetic

Whilst the request was made for all participants to be fluent in English, no stipulation was given for such fluency to be of native origin, thus meaning participants could be fluent speakers of English from either a native or non-native perspective. In so doing, it would permit the comparison of responses from both groups, and further cross-comparison between participants with fluency from a different second language, thus giving a comparative subset equal to the amount of represented languages (**n**), as shown;

English (native) < > Language ***n*** (second)
Language ***n*** (native) < > English (second)

The purpose of comparison was again to establish how each participant would compare in their emotional and aesthetic scoring, but furthermore to analyse subsequent differences or similarities between native and fluent non-native English speakers. This was further noted following a personal interview with one of the participants who stated that second languages are likely to be more aesthetically pleasing simply as a result of their being different to one's native tongue (Mancisidor, 2014). As sufficient results were received to allow for such an analysis, the hypothesis was tested and proved correct in this context (see figure 16). However, should the languages be further subcategorised into groups relating to their respective use (family, work, social interaction etc.), I would conjecture that such results would further indentify emotional conditioning. Much like the paradigm associated with the first experiment, it is likely that the same difficulty in separating personal connotations from the language would occur. For example, native languages are often used for family communication during childhood, and thus are likely to inherently induce positive emotion, whereas those who adopt their second language for career purposes may not find such an aesthetically positive response.
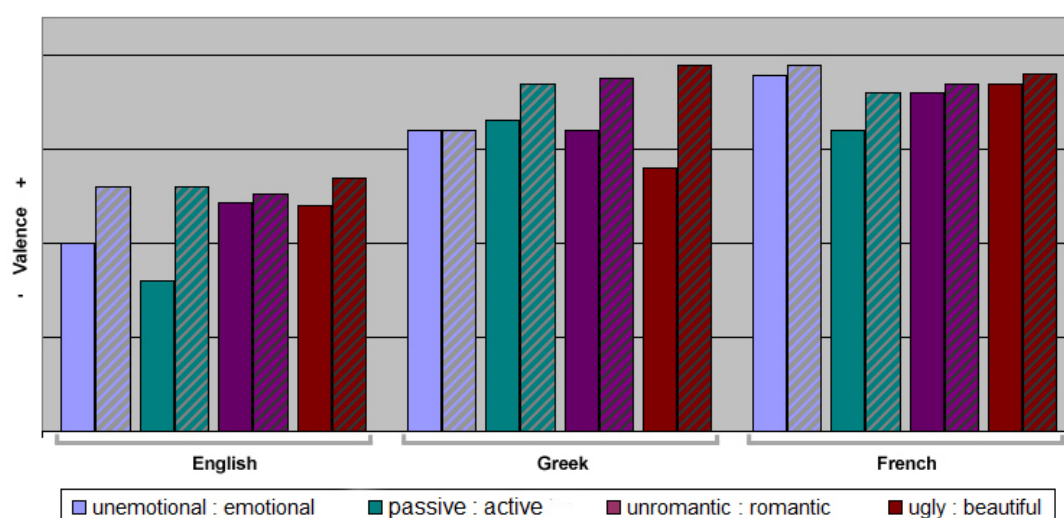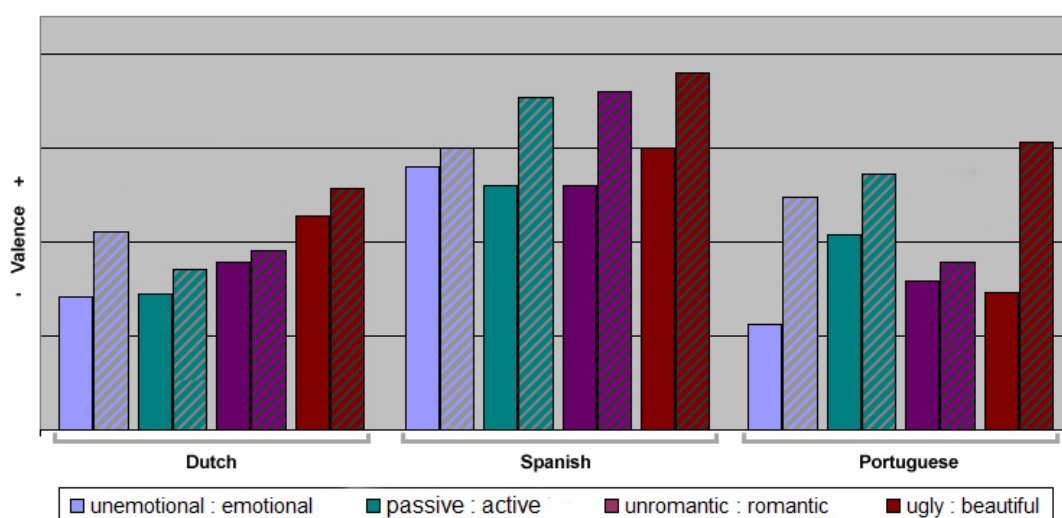
Figure 16 - Responses of native speakers scores against non-native (second) language speakers. Adjacent shaded areas show non-native responses [20].

As all participants were fluent in English, either in a native or non-native capacity, it seemed prudent to further extend the analysis from the results shown above by comparing the responses of native English speakers against those using English as their second language. As anticipated by the previous results, non-native users rated every aspect of English above their own native language. One exception being the responses from native Greek speakers who, as anticipated by previous results, rated the affection of their native tongue above that of English. The other, again somewhat anticipated, was that of native French speakers who rated *every* aspect of their language above English.

---

[20] Please note that insufficient results were obtained from non-native Russian speakers for inclusion into this analysis.
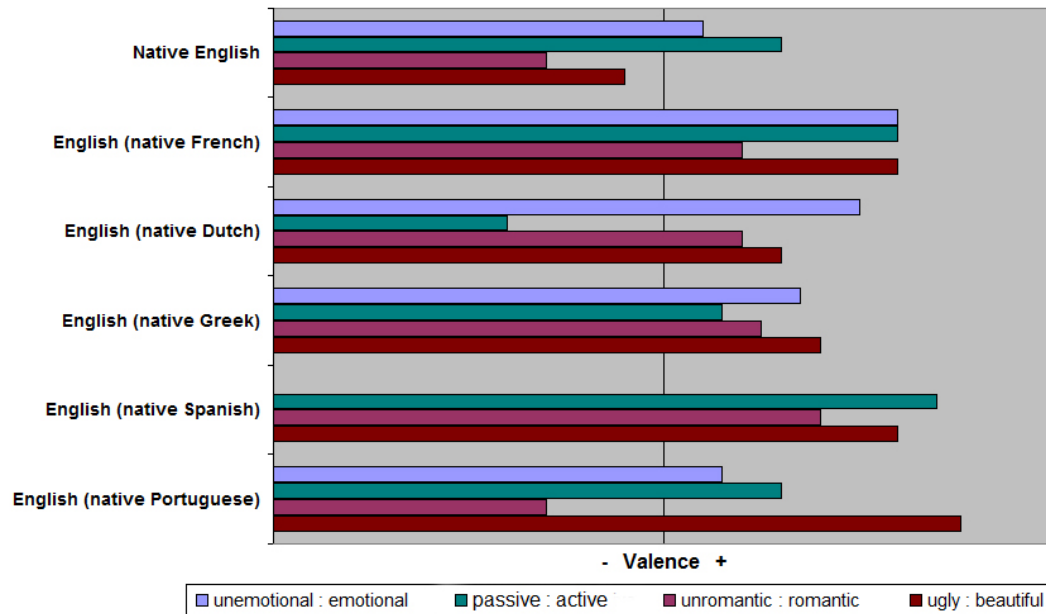
Figure 17 - Scoring of English shown against non-native English speakers [20].

The subsequent part of the survey was written to establish similarities and differences between language phonemics, and the emotional impact they have on the written or spoken word. The process began first by asking participants to note words, of *any* language, which they found to be either beautiful or ugly. As bilinguals are likely to encounter words and phrases from other languages, by permitting the use of any language, thus giving a wider range from which to respond, the results could only benefit from any potential diversity. The results showed that the native language was the main focus for the *negative* words, whilst the second language being the primary focus for *positive* responses. Those which were taken from a language other than the participants known languages were evenly distributed between both positive and negative results (figure 18).

NATIVE LANGUAGE

negative
25%

positive
10%

12%
negative

UNSPOKEN LANGUAGE

15%
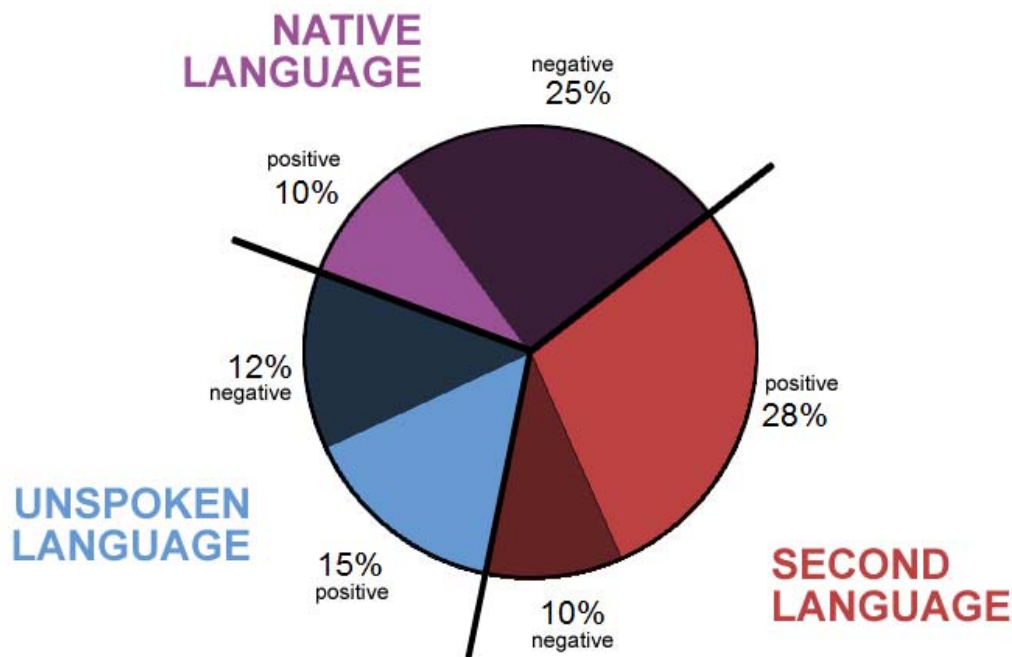positive

positive
28%

10%
negative

SECOND LANGUAGE

Figure 18 - Rating of emotionality in words chosen by bilingual participants

The results further evidence the aforementioned hypothesis that a second language is favoured in terms of emotionality; Dutch responses more so than any other, with native responses contributing to over 90% of negative word list. Although once again, native Greek were an exception to this evidence, as were both native and non-native French speakers. Both avoided using English for positive responses, and instead used words of either native or unspoken origin.

However, whilst there appears to be an overall emotional preference for non-native language, the question still remained as it did for the aforementioned monolingual experiment. Are there any patterns of phonetic preference in either positive or negative responses between languages? And could any similarities be found between those favoured by bilinguals and monolinguals? To this end, a further analysis using the same criteria as the monolingual surveys was then carried out for the above responses. If there is any possibility for there being a set of phonaesthetic fundamentals governing all languages, this would initially be evident in the results of such an analysis.
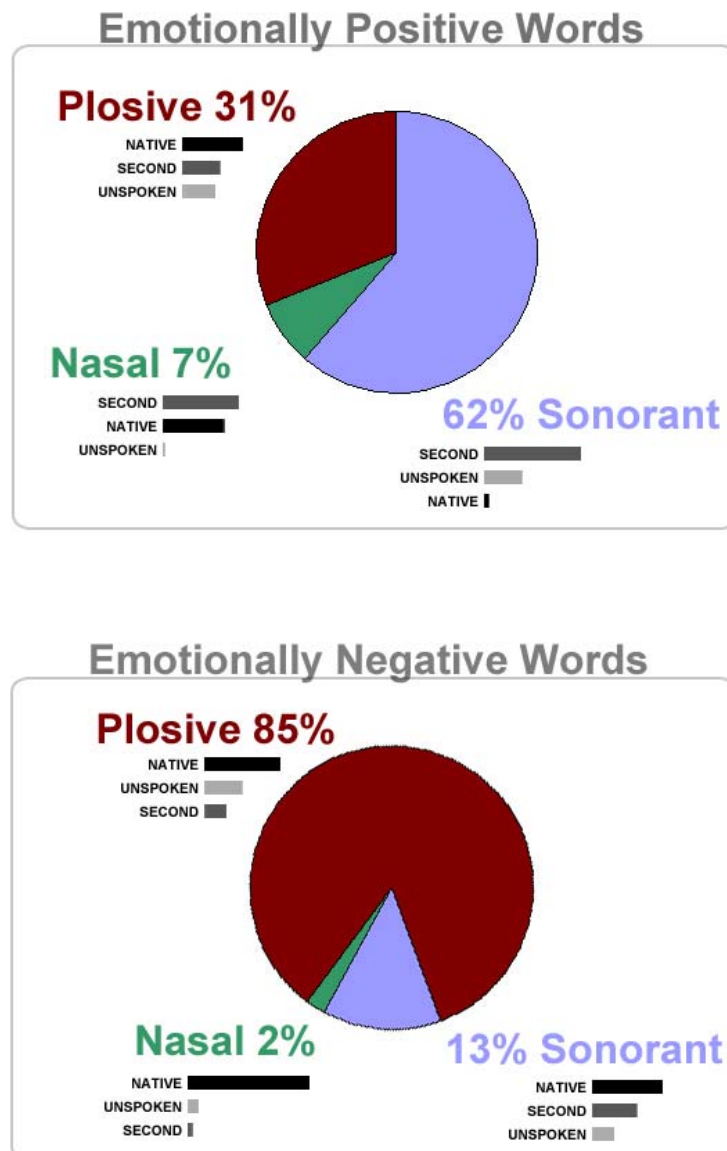
Figure 19 - Results of a phonemic analysis of emotional words
chosen by bilingual participants

Whilst it cannot be concluded from these initial results whether they corroborate with those from the first survey, they do further evidence the favoured second language hypothesis. The bilingual's second language constituted 53% of the emotionally positive words, whilst unspoken and native words occupied 28% and 19% respectively. Conversely, the negative emotional words consisted of 53% native words, 26% unspoken words, and 21% of words from the second language. To sum, the second language is more than twice as

likely to be recognised as a positively emotional lexicon, in the same way that the native is recognised as a negatively emotional language.

For the final part of the survey, it was decided to isolate phonemes rather than use entire words, and again analyse the responses from participants rating them on aesthetic favourability. If a universal aesthetic did indeed exist, and that each phoneme is attributed with specific emotional responses as demonstrated by Whissell's research, then the same results would be shown for native English bilinguals. However, assuming this to be the case, would there also be a similar pattern shown in the results of the non-native English participants?

As consonants can easily be misinterpreted in other languages, this test concentrated on vowels, being more phonetically uniform across the languages represented in the present study. Pairs of vowels were given, of which one was to be chosen in preference over the other, and of which best represented emotional positivity [21]. For example, an open front vowel (a, apple) was placed alongside an open back vowel (ɑ, art), whilst others were adjacent to mid-position vowels, thus giving a balance of all possible vowel positions. The words were also paired into what are referred to as short and long vowels [22]. The purpose of so doing was to isolate the vowel phonemes to ascertain whether the vowel length, or vowel position contributed to the emotional preference; something of which would not have been uniformly possible using consonant phonemes. The results are shown in the following diagram (figure 20).

---

[21] Whilst participants were required to equate each vowel phoneme to either a negative or positive category, arousal was not specified (passive, neutral, active) due to the absence of example words. The motivation for doing so was to allow the participant to think of words by themselves, in any language they so desired.

[22] The terms short and long vowel relate to the duration of the pronounced phoneme, or what is also referred to as vowel length. Words such as "apple" and "art", or "import" and "island" demonstrate the difference between vowel length (short and long respectively). Both pairs also demonstrate the difference between front and back vowels, coinciding with the short and long vowels.
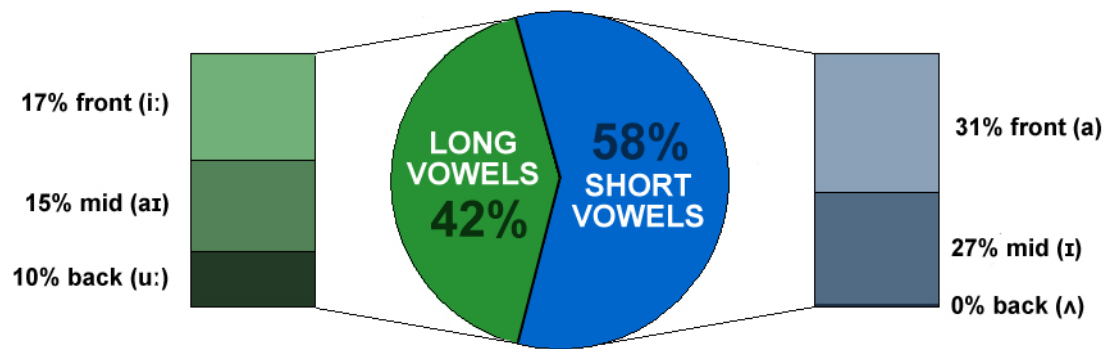
Figure 20 - Results of vowel length and position test on bilingual participants

Overall, the results showed back vowels to be the least emotionally communicative; less than 24% of the long vowels were represented by the back vowel, whilst the short back vowel was not chosen by any of the participants. As previously stated, reasons for rejecting the back vowel as emotionally positive have already been the focus of conjecture by Fónagy, among others. Whilst research has shown that back vowels are more frequently used in expressive texts when referring to slow and heavy movement (Fónagy, 1961: 194), in other cases they have been utilised in portraying dark colours, whilst front vowels have been associated with light colours (Jackobson and Waugh, 1979: 188). Fónagy's aforementioned poetic research also shows negative moods to be associated with back vowels, where they are used favourably to describe aggression and negativity. A further explanation for this tendency may also be associated with the aforementioned metaphor of negative speech patterns being characteristic of nausea and coughing. Whilst Fónagy had directed this hypothesis to the observations of plosive phonemes, in the same way, back vowels move the same speech organs, thus making the same metaphor plausible in this respect.

Crystal conducted similar tests that also allowed for vowel position analysis, of which showed there to be a greater concentration of front vowels in aesthetically pleasing words (more so than back or mid-position vowels) than found in "everyday conversation" (Crystal, 1995:11). The same results also

indicated a preference for polysyllabic words to contain vowels in either a front or mid position, or to move in a front-to-back direction during syllabic pronunciation (see diagram 21), which may suggest why Tolkien found an affinity with the words "cellar door". But again, these results were limited to words chosen by native English speakers under conditions which would not allow for semantic ignorance, and thus by targeting non-native English bilinguals, a better indication of a universal aesthetic could be indicated.

camellia (kəˈmɛlɪə)   **cellar door** (sɛlə dɔː)

mid - front - front - mid   front - mid - back

Figure 21- Examples of aesthetically-pleasing polysyllabic words
using David Crystal's phonetic algorithm

However, when these results were placed alongside those from Whissell's experiments, there was little evidence to support a universal phonetic aesthetic. Both the front (a) and back (ʌ) short vowels were shown to be passive, but valence was not specified. So whilst the front vowel may corroborate a positive passive response, the back vowel was not chosen by any participants, thus opposing the results of Whissell's tests. The short mid vowel (ɪ), again scoring above average in the above responses, was shown in Whissell's results to be both positively *and* negatively active. Again, whilst parallels can be drawn between the positive responses, the ambiguity discredits any substantial accuracy between results. Finally, Whissell's front (ʌɪ) and back (iː) long vowels were shown to be unanimously positively, neither of which were chosen over the short vowel counterparts in the above results. Suffice to say that whilst native speakers may agree with the emotional representation of selected vowels, a comparable cross-cultural paradigm does not exist.

Furthermore, by separating the results into the speakers respective native language, little could be found in terms of preference for exclusive vowel location, with the exception of both native and non-native French bilinguals, of whom chose the back vowels, in long *and* short cases, over both mid and front positions. Setting aside the aforementioned metaphors, an observable characteristic of back vowels is the increased phonemic duration which, following the information gleaned from the French bilingual interviews, would explain why it attracted the given response. But how do such phonemes become a characteristic of what is considered a positively emotional and romantic language?

Looking back to the aforementioned emotional linguistic spectrum, at the far end of the scale we recognise that music encodes emotion in a greater concentration than the spoken word. Although it isn't the intention of the present paper to address the studies which have shown this to be the case, research has shown how phonetic characteristics can demonstrate musical qualities under certain conditions. As a result, this may suggest the languages ability to encode emotion to a greater extent of concentration. The studies responsible claim to have isolated the elements of speech that perceptually transform words into music, or what has since been referred to as the *Speech to Song Transformation* (S2ST). One of the characteristics found to exhibit features that would determine such a transformation is that of pitch, or *lexical tone*. Musical vocalisation relys on a system of discrete pitch values that in turn produce an interval structure that adheres to a musical scale. Conversely, speech does not rely on such scales for successful communication, and the pitch differences are evaluated in the context of the discourse (Rathcke et al, 2014:1492). Therefore, where an utterance of speech remains closer to discrete pitch values, it further facilitates the encoding of emotion by demonstrating the same musical qualities. So, given the limited variance of the F_0, and the lengthened vowel duration (as characterised in part by the use of back vowels), it is plausible to say that the French language also exhibits musical qualities in this respect. As a result, this suggests that phonetic attributes can be perceived differently under specific conditions, and that it is

the context in which the phonemes are used in the French language that determines their emotional perception.

Conclusion: The Aesthetic Alternative

Considering the combined information presented in this paper, suffice to say that a considerable amount of research has been targeted towards unifying the individual elements of language; the way in which the spoken and written word is delivered, the means by which it is received or interpreted, and the modalities and cues that attribute to the exchange and processing of communication. These areas have received recognition from a variety of academic disciplines, and from the works of authors such as those cited in the present paper, we are able to build upon the theories presented by Saussure, Jakobson, Chomsky and Sapir to gain a better understanding of how the artistic aspects of language can be further recognised, appreciated, and used to greater potential in communication.

Whilst the concept of there being a set of universal aesthetic phonemes remains increasingly unlikely, it is plausible to suggest that a combination of theories and hypotheses may increase the likelihood of achieving a close approximation within the context of a given language or culture. From the results shown in the constructed language tests, there are undoubtedly phonemic characteristics that will permit the further encoding of emotion to better than chance levels, thus resulting in what Crystal refers to as an ideal phonaesthetic language. For example, the aforementioned "cellar door". Whilst the phrase makes use of a back vowel, which in isolation may not be aesthetically pleasing, the phoneme is in a word-final position (assimilating the /r/), and thus becomes pleasing as a result - perhaps as a result of the word demonstrating a uniform decrease of F_0, at what is considered appropriately aesthetic for English phraseology. Whilst conjecture remains, it is plausible to suggest that further research of phoneme positions in this

context may yield further results in favour of an algorithm of which will also facilitate the creation of the aforementioned ideal phonaesthetic language.

Returning to the extension of Saussure's hypothesis of concept and sound image (figure 8), we now find that the same extension becomes subject to further criteria as indicated by the present paper. Whilst the universal connection between the concept and sound image is not without merit, in the context of emotional recognition, both the concept and the sound image hold differentiating levels of concentration when applied to the proposed extension.

In accordance with the results in the previous tests, the emotional perception that follows the sound image becomes subcategorised into phonetic and cognitive recognition. Then, depending on whether or not the sound image is itself recognised, the ratio of concentration becomes biased towards one of the two levels of recognition, as evidenced in the above monolingual tests (figures 10, 11 & 12). That is to say, by diverting the focus away from the implied concept, the level will be biased towards the phonetic recognition at a rate comparable to the reduced level of cognitive recognition (figures 4a & 4b). That is to say, an English monolingual will have a far greater concentration of phonetic recognition as a result of hearing Chinese or Japanese, due to the phonemic structure compared to the native language, but not as great as French or Spanish, whose phonemics structure reflects a greater amount of the English phonology. However, an unrecognised (constructed) word without a given concept, or what Sapir referred to as a phonetically impossible word, will increase the ratio further in favour of the phonetic recognition. Although the same would inevitably incur an element of cognitive recognition due to the analysis of phonemic patterns and subsequent comparison to the participants native language, which would arguably increase with the use of bilingual or multilingual participants, again given their increased reference of phonemic characteristics. The following diagram visualises the suggestions shown above, again based on the Saussurean model.
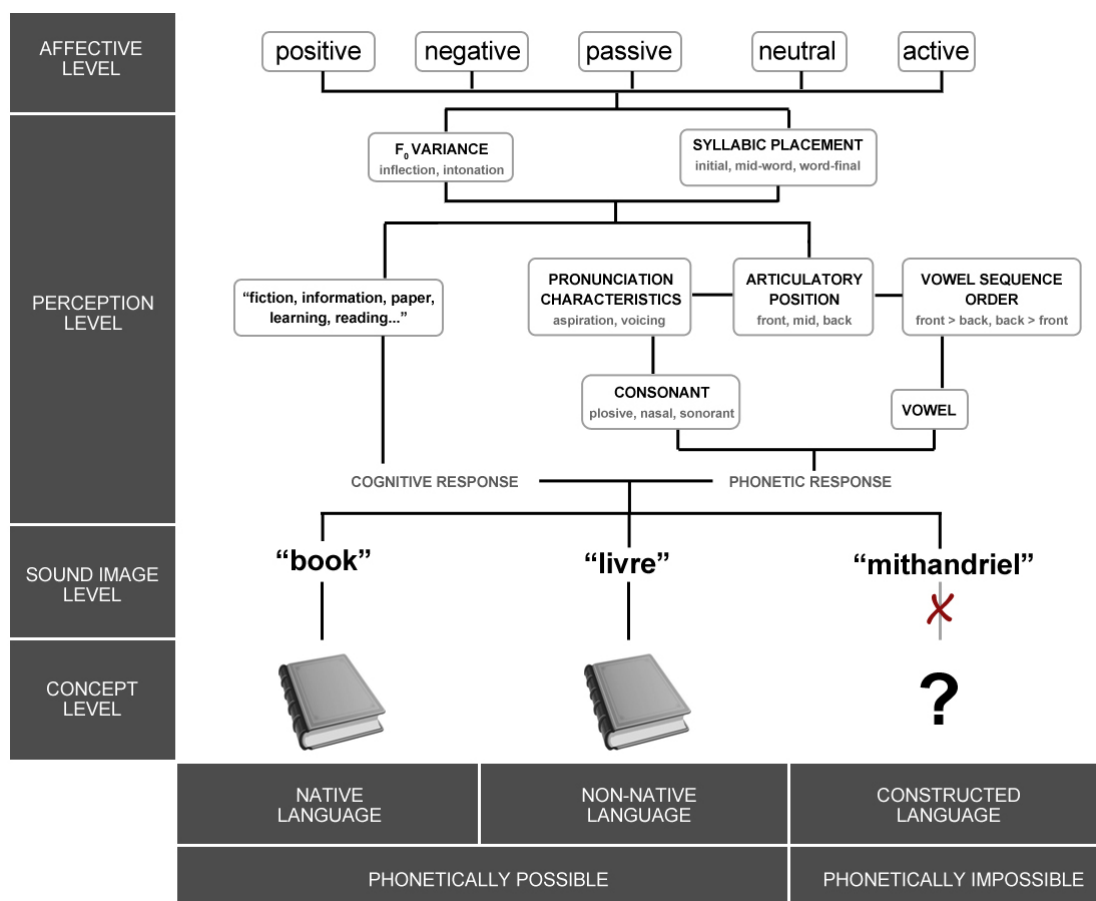
Figure 23- The collective extension of the Saussurean model incorporating an emotional component

Whilst this model covers the hypothesis of how emotionality is perceived in respect to the spoken and written word, there remains a multitude of other factors pertaining to non-lexical emotional stimuli, which are likely to further influence the levels and bias of concentration. For example, how will such a model allow for the difference in alternating pitch between gender? How will it be appropriated for the addition of visual stimuli? And to what extent do the above bias the concentration between cognitive and phonetic recognition? Not only must we allow for the context of the dialogue and the appropriate sound imaging, but at what point do visual elements determine a diversion of attention away from the audible stimuli? For example, within the context of a film or TV series, how does the aesthetic beauty of a portrayed character, or the detail of the environment influence such attention? How does this alter

the paradigm of native and constructed language? Does a personal relationship with the speaker or the environment alter the perceived levels of emotion? Perhaps it is possible to arrive at appropriate algorithms deduced from the participants native language, as done so for the above model. Or maybe Saussure's model will be sufficiently rewritten beyond recognisable merit. Suffice to say that whilst emotional research within the field of aesthetics is becoming a recognised branch of linguistics and neuroscience, we are nonetheless scratching, an albeit exciting, surface.

Acknowledgements

Cited Works & References

Carroll, L. *Alice's Adventures in Wonderland and Through the Looking-Glass*, Penguin (2010)

Crystal, D. *Phonaesthetically Speaking* (English Today 42, Vol. 11, No. 2), Cambridge University Press (1995)

Crystal, D. *Still Finicking with Trifles, Largely*, Otto Jespersen Memorial Lecture (1995)

Erkkilä, A. Various - Personal communication via social media site relating to aspects of Finnish phonology (2014)

Falk, JS. *Otto Jespersen, Leonard Bloomfield, and American Structural Linguistics*, Language 68(3):465-491 (1992)

Fonagy, I. *Why Iconicity? In Form Miming Meaning: Inconicity in language and literature* (1999)

Fonagy, I. *Paralinguistic universals and preconceptual thinking in language* (1991)

Heiser, M., Iacoboni, M., Fumiko. M., Marcus, J. & Mazziotta JC. *The essential role of Broca's area in imitation*, European Journal of Neuroscience, Vol. 17, Issue 5, pp 1123–1128 (2003)

Ilie, G. & Forde-Thompson, W. *A comparison of acoustic cues in speech for three dimensions of affect* (2006)

Jakobson, R. & Waugh, L. *The Sound Shape of Language,* Indiana University Press (1979)

Koch, C. *Why your brain has a Jennifer Aniston cell* , pub. 22 June 2005, http://www.newscientist.com/article/dn7567-why-your-brain-has-a-jennifer-aniston-cell.html#.U_pc3Gemf4Y [online - accessed 24th May 2014]

Lasbareilles, C-L. Personal Interview for French and European culture and language survey results (October 2014)

Lasbareilles, C. Personal Interview for French culture language survey results (October 2014)

Lauer, G. *Spiegelneuronen. Uber den Grund des Wohlgefallens an der Nachahmung*, Karl Eibl/KatjaMellmann/Rudiger Zymner (ed.), Im Rucken der Kulturen, Paderborn, pp. 137-165 (2007)

Leech, G. *A linguistic guide to English poetry*, Longman Press (1991)

Loui, P., Bachorik, JP., Li, HC. & Schlaug, G. *Effects of voice on emotional arousal*, Frontiers in Psychology, Vol. 4, Article 675 (2013)

Maciamo, A. *Number of Phonemes by language in Europe* http://www.eupedia.com/forum/threads/29850-Number-of-phonemes-(vowels-consonants)-by-language-in-Europe [online - accessed 5th August 2014]

McCaffery, S. *Sound Poetry - A Survey* http://umintermediai501.blogspot.co.uk/2008/01/sound-poetry-survey-steve-mccaffery.html [online - accessed 15th July 2014]

Mancisidor, G. Personal interview for language cultures in Spain and Portugal (May 2013)

Majid , A. *Current Emotional research in the language sciences*, Emotion Review, Vol. 4, No. 4 (2012)

Ovadija, M. D*ramaturgy of Sound in the Avant-garde and Postdramatic Theatre*, McGill-Queen's University Press (2013)

Perlovsky, L. *Language and cognition - joint acquisition, dual hierarchy and emotional prosody*, Frontiers in Behaviour Neuroscience, September 2013, Vol. 7, Article 123 (2013)

Perlovsky, L. *Language and emotions: Emotional Sapir-Whorf hypothesis,* Neural Networks Ed. 22 Special Issue (2009)

Perlovsky, L. *Musical emotions: Functions, origins, evolution,* Systemics, Cybernetics and Informatics, Vol. 11, No. 9 (2013)

Podhorodecka, J. *Is Iamatyave a linguistic heresy?* (extract from Insistent Images), John Benjamins Publishing Company (2007)

Pütz, M; Verspoor, Marjolyn, eds. *Explorations in linguistic relativity*, John Benjamins Publishing Company (2000)

Rathcke, T., **Falk, S.** & **Dalla Bella, S**. When Speech Sounds Like Music, Journal of Experimental Psychology: Human Perception and Performance 2014, Vol. 40, No. 4 (pp. 1491-1506)

Reuven, T. *How Do Sound Patterns Know They Are Expressive? The Poetic Mode of Speech Perception* (1992)

Salgaro, M. *The Text as a Manual, Some Reflections on the Concept of Language from a Neuroaesthetic Perspective* (2009)

Sapir, E. *Selected Writings of Edward Sapir*, University of California Press (1963)

Saussure, F. *Course in General Linguistics* (1916 ed.), Columbia University Press (2011)

Scobie, S. *I Dreamed I saw Hugo Ball: bpNichol, Dada and Sound Poetry*, boundary 2, Vol. 3, No. 1 (1974)

Suomi, K., Toivanen, J. & Ylitalo, R. *Fonetiikan ja suomen äänneopin perusteet.* Helsinki: Gaudeamus Kirja (2006)

Taylor, IK. & Taylor, MM. *Another look at phonetic symbolism.* Psychol Bull., Dec; 64(6):413-27 (1965)

Whissell, C. *Sounds and Feelings*, unpublished material obtained via email correspondence (2014)

Wiese, H. *Numbers, Language, and the Human Mind* (2004)