

Emotional Prosody in Interjections: A case of Non-arbitrariness in Language

Åsa Abelin¹

*Department of Philosophy, Linguistics and Theory of Science
University of Gothenburg*

Emotional prosody shows a connection between meaning and expression, and constitutes a special case of non-arbitrariness in language. Emotional interjections are learned at an early age, and also have a biological basis for their expression. It is therefore possible that emotional prosody of interjections is part of the phonological and semantic representation in the mental lexicon, and hence can be expected to influence visual lexical decisions. If confirmed, this would have some bearing on the debate concerning the relations between ‘linguistic’ and ‘affective’ prosody. A priming experiment was performed on Swedish interjections with two emotional meanings: HAPPINESS (positive words) and DISGUST (negative words). The main question was whether emotional prosody primes written interjections with emotional content, through cross-modal priming, and the chosen method was to elicit lexical decisions in a cross-modal priming task and in isolation. The results show that there was an effect of priming, and that the effect was significantly greater for HAPPINESS words (and HAPPINESS prosody) than for DISGUST words (and DISGUST prosody). For individual words, there was a positive correlation between a high priming effect for the corresponding emotion and the degree of correct interpretations of emotional primes. Furthermore, there was a tendency for high-frequency words to be primed more than low-frequency words, when the emotion of the prosody was matched. There was no such effect for high-frequency words when the emotion of the prosody was mismatched. There was also a tendency to a negative correlation between degree of correct interpretations of emotional primes and high priming effect, when the prosody was mismatched. We interpret these results to mean that it is problematic to regard emotional prosody as non-linguistic and disconnected from the lexicon, since there was a gradual connection between spoken emotional prosody, written emotional interjections, and lexical frequency of interjections.

1. Introduction

1.1. Non-arbitrariness and emotional prosody

Linguistic signs can be both non-arbitrary (motivated) and conventional. Furthermore, the three basic semiotic relations—contiguity, similarity, and convention—are not mutually exclusive, and can be seen as determining the signifier–signified relation to various degrees (Ahlner & Zlatev, 2011). The semiotic relation contiguity can also be termed indexical, and the relation similarity can be termed iconic; both of these are non-arbitrary. This article focuses on the indexical relation between signifier and signified in emotional prosody.

The emotional prosody which is present in emotional interjections is a special case of non-arbitrariness in language. The non-arbitrary relation between prosodic expression and the meaning of emotional interjections is mainly indexical, since there is a spatiotemporal, or even causal, relation between content (emotion) and expression, when considering the bodily aspect of emotion. Phonaesthemes and emotional prosody share a number of traits: they are non-arbitrary (indexical), express emotional meaning, and use the so-called frequency code (Ohala, 1994). The frequency code refers to the use of acoustic signals with low frequencies to convey that the signaller is large, potent, and

¹ Email: asa.abelin@ling.gu.se

dominant, and the use of signals with high frequencies to convey the opposite (cf. Johansson & Zlatev, this volume).

Lexical gangs (Pierrehumbert, 2010) are sets of words with shared phonological and semantic properties that influence morphological productivity. Phonaesthemes are a special case of lexical gangs (Abelin, 1999; Bergen, 2004). Emotional prosodic morphemes could constitute another special case of expressions with gang behaviour. Emotional interjections are learned at an early age, and it is possible that emotional prosody is part of the phonological and semantic representation in the mental lexicon for these interjections (cf. Nygaard & Queen, 2008; Abelin, 2013). If this is the case, emotional prosody can be expected to influence written emotional words. The question is thus whether nonsense words spoken with different emotional prosody can influence lexical decisions for visually presented words with matched or mismatched meanings, and whether it is really fruitful to distinguish between ‘linguistic’ and ‘affective’ prosody.

The acoustic signal of prosody, especially intonation, is a bearer of many types of information: emotional, attitudinal, syntactic, lexical, interactive, informational, dialectal, and personal. Speaker characteristics are clearly non-linguistic, but it is not self-evident that emotional or attitudinal information would be so. Emotional prosody is traditionally often described as non-linguistic or extra-linguistic; that is, as not really belonging to the linguistic system (e.g. Zsiga, 2013: 68). Linguistic prosody would then be reserved for conveying lexical contrast, for example by word accent or tones, or being connected with grammatical function such as contrastive focus or questioning. However, within a cognitive theory of language there could be room for less clear-cut boundaries between so-called linguistic prosody and emotional prosody. Seddoh (2002) also points out that emotions and attitudes may be signalled by words or phrases, without words and phrases being considered non-linguistic. Stankiewicz (1964) has shown that stress shift, code switching, substitution of grammatical constructions, and phonological processes are also used to signal emotions.

1.2. Morphological priming

Morphological priming effects have been shown to be stronger than purely semantic or purely phonological priming (see, e.g., Bergen [2004] for a review). This has been used to argue for the psychological reality of morphemes. Bergen (2004) found that phonaesthetic priming was significantly stronger than purely phonological or semantic priming, and thus phonaesthemes seem to be psychologically real in the same way that morphemes are. Similar results for Swedish were found by Abelin (2012). The question remains whether emotional prosody has any priming effect on reactions to written emotional interjections. If there are priming effects that are at least as strong as morphemic effects, this might be an argument for emotional prosody or ‘emoproemes’ being a linguistic unit at or above the morpheme level. An emoproeme would be a unit of emotional meaning (categorical or dimensional) and a prosodic profile, and furthermore non-arbitrary in nature, just like phonaesthemes. A sketchy example of an emoproeme could be the unit of HAPPY emotional meaning and a prosodic profile with an F0 peak at the stressed syllable,² large F0 range, short duration, and strong intensity. It could be argued that there is great variability and relativity in the acoustic manifestation of the prosodic profile, but there is no evidence that this variability should be larger than for vowels and consonants. Emoproemes would be obligatorily specified in interjections.

² The placement of the emotional F0 peak not being independent of where word stress is placed would in itself be an argument for emotional prosody being connected with the rest of grammar.

1.3. Cross-modal priming

Cross-modal priming is a well-known phenomenon in different areas of language, but research on cross-modal evaluative priming has been limited. Cross-modal priming means that a stimulus in one modality will facilitate the recognition of a related stimulus in another modality, for example between hearing and reading or between hearing and looking at faces. Scherer and Larsen (2011) gave an overview of earlier research and performed a study on cross-modal evaluative priming. They studied how positive sounds (e.g. sounds from a seagull or a robin) and negative sounds (e.g. vomiting or screaming) prime written evaluative words like sunny or luck (positive) and murder or agony (negative). They found that negative auditory primes have a greater effect on visual word processing than do positive auditory stimuli. They also concluded that cross-modal priming research needed to be performed on qualities of the human voice, for example, pitch. Schirmer, Kotz and Friederici (2002) investigated the influence of emotional prosody on the processing of visually presented positive and negative target words in a cross-modal priming study using the event-related potential (ERP) technique. The results indicate that emotional prosody modulates word processing. There have also been claims that processing differences exist between positive and negative emotions, whereby positive emotions give a processing advantage in different contexts (cf., e.g., Becker, 2012; Nobata, Hakoda, & Ninose, 2010). The present study of positive vs. negative emotions aimed to discover the effects of types (dimensions) of emotion. Thus, the entities of relevance for the present analysis were emotion, emotional type (positive/negative), and individual emotional interjections.

1.4. Research questions

Four studies are presented. Experiments 1 and 2 aimed to examine whether emotional prosody primes reactions to written interjections (Experiment 1), and whether this priming differs from lexical decision in isolation (Experiment 2). Experiments 1 and 2 also aimed to examine the difference between positive and negative emotions (for matching primes and targets), and to investigate whether there is a correlation between priming facilitation (Experiments 1 and 2) and a measure of accuracy in interpretation of the primes (Experiment 3). Word frequencies were extracted from a Swedish corpus (Study 4) since the frequencies of words used as targets could be expected to affect reaction times to these targets. The relation of priming facilitation and accuracy in interpretation of the primes in relation to word frequency was inspected.

2. Experiment 1: Method

The method used in Experiment 1 was a priming task in which the participants were presented with sequential pairs of words. There can be a processing advantage for the second word (a facilitatory priming effect) if the words have some similarity (e.g. semantic or phonological). In the present experiment, the question was whether spoken non-words with two different types of emotional prosodies – HAPPINESS and DISGUST – would give a priming effect in written emotional interjections. For example, the experiment might test whether the non-word **fuj*³ spoken with happy prosody would produce a facilitatory priming effect in the written word *hurra* ('hurrah').

³ Note that the grapheme “j” in Swedish corresponds to the (semi-)consonant /j/, approximately corresponding to the English grapheme “y” as in *you*, not to “j” as in *John*.

2.1. Material

Part of the stimuli were in the semantic condition, meaning that the prime and the target words shared a semantic feature but no expressive feature. For example, the prime was nonsense (when one considers the sequence of vowels and consonants), but with emotional prosody expressing HAPPINESS or DISGUST. The target was a written emotional interjection. The targets were written Swedish interjections with meanings related to HAPPINESS or DISGUST; five of them expressed HAPPINESS and three expressed DISGUST. These were repeated twice during the task (see Table 1).

Part of the stimuli were in the ‘baseline’ or mismatched condition, meaning that prime and target were unrelated both in form and in meaning. Examples are the non-word **fuj* with HAPPINESS prosody and the written interjection *usch* (‘ugh’) which expresses DISGUST. The non-words with prosody unrelated to the target words in emotional content, were not expected to give priming effects. These words were partly the same as the ones used by Abelin (2011) as non-word fillers. The primes and the matched targets are shown in Table 1. The items were repeated in the tasks.

Table 1. Primes: spoken non-words with different emotional prosodies; targets: written emotional interjections

Prosody	Spoken non-word prime	Written target	English gloss
HAPPINESS	<i>fuj</i>	<i>hurra</i>	‘hurrah’
HAPPINESS	<i>jeffe</i>	<i>yes</i>	‘yes’
HAPPINESS	<i>naha</i>	<i>tjoho</i>	‘yippee’
HAPPINESS	<i>ogg</i>	<i>ja</i>	‘yes’
HAPPINESS	<i>ply</i>	<i>jippi</i>	‘yippee’
DISGUST	<i>hule</i>	<i>usch</i>	‘ugh’
DISGUST	<i>tjahe</i>	<i>blä</i>	‘ugh’
DISGUST	<i>åbb</i>	<i>fy</i>	‘ugh’

There were also non-word fillers among the targets. In total there were 9 test targets (repeated twice) and 18 non-word filler targets. The effect of the stimuli being repeated twice did not give any ‘internal’ priming effect; that is, no facilitation from the first occurrence of an interjection to the second one (cf. Waern Kyrck, 2013).

2.2. Participants and procedure

The participant groups in Experiments 1 to 3 were non-overlapping; that is, no individual participated in more than one experiment. The participants in Experiment 1 were 26 native speakers of Swedish (mean age: 22 years, 24 women), all of whom performed a lexical decision task with priming. Thirty-six prime stimuli of different emotional prosodies appeared until they self-terminated, for approximately 462–652 msec for non-words with HAPPINESS prosody and 690–1037 msec for non-words with DISGUST prosody. The difference in duration is inevitable, but did not affect the results; for a comparison, see discussion by Abelin (2013.) Target stimuli appeared 300 msec later, for 1000 msec or until the subject responded to them. The maximal response time was 1000 msec. The stimuli were in fixed order for all subjects. The participants were asked to decide as quickly and as accurately as possible whether or not the target word was a real Swedish word, and to respond by pressing one of two keys on the keyboard of a MacBook Pro. The reaction time program PsyScope X (version B57) was used for the

experiment. All tests took place in a silent room.

3. Experiment 2: Method

The method used in Experiment 2 was a lexical decision task. The target words (written Swedish interjections) of the priming test were here tested in isolation, in ordinary lexical decision experiments, in order to check the effect of priming in comparison with no priming. Experiment 2 is also described and reported in more detail by Waern Kyrck (2013).

3.1. Material

The stimuli were the same as in Experiment 1, shown in Table 1, apart from the primes. Thus there were 9 test targets (repeated twice) and 18 non-word filler targets. The targets were written interjections, five expressing HAPPINESS and three expressing DISGUST. These were repeated twice during the task.

3.2. Participants and procedure

The participants were 13 native speakers of Swedish (mean age: 20 years, 9 women) who performed a lexical decision task with written interjections in isolation. Target stimuli appeared for 1000 msec or until the participant responded to them, with the maximal response time being 1000 msec. The interjections were the same targets as in Experiment 1, presented in the same order for all participants. The participants were asked to decide as quickly and as accurately as possible whether or not the target word was a real Swedish word, and to respond by pressing one of two keys on the keyboard of a MacBook Pro. The reaction time program PsyScope X (version B57) was used for the experiment. All tests took place in a silent room.

3.3. Types of analyses

The reaction times for primed targets, in matched and mismatched conditions, were compared with the reaction times for non-primed targets in the lexical decisions in isolation. Comparisons of reaction times between HAPPINESS and DISGUST emotions were made. Furthermore, the differences in reaction times for primed and unprimed targets were correlated with success of interpretation of nonsense words with different prosodies (see Experiment 3) and with a frequency count of the target words (see Study 4).

4. Experiments 1 and 2: Results

The results from Experiments 1 and 2 are here presented in combination, since the reason for performing the two experiments was to study the difference in reaction times between them. A comparison of the mean reaction time (RT) in the priming experiment with that in the lexical decision experiment with written interjections in isolation showed a priming effect for both matched and mismatched conditions. (However, the priming effect varied for individual words; see Figure 2.) The comparison also showed that the difference between primed and isolated targets was larger for happy than for disgusted interjections; that is, the primed happy interjections had shorter RT (see Figure 1). The priming effect was significantly larger for HAPPINESS words than for DISGUST words.

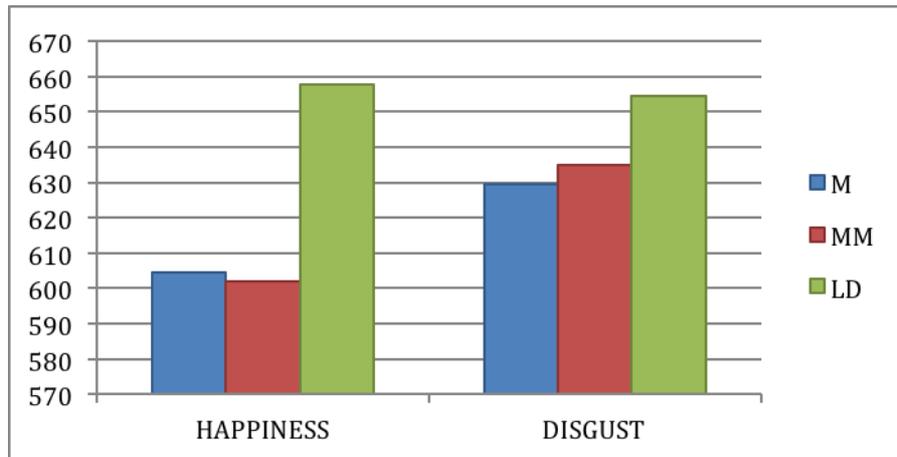


Figure 1. Priming effect in both matched and mismatched conditions as compared with lexical decisions without priming. Mean reaction times for matched, mismatched, and lexical decisions in isolation, for the emotions HAPPINESS and DISGUST

T-tests showed significant differences between HAPPINESS lexical decision (LD) and HAPPINESS matched (M) (two-tailed, $p = .0090$) and between HAPPINESS LD and HAPPINESS mis-matched (MM) (two-tailed, $p = .0026$). This means that, in general, HAPPINESS words were more easily primed than disgust words by the emotional prosodies. There were no significant differences between HAPPINESS and DISGUST in LD, between DISGUST LD and DISGUST M, or between DISGUST LD and DISGUST MM.

4.1. RT differences for individual interjections

Figure 2 shows the RT difference between the conditions of priming and isolated lexical decisions, for individual words. Individual words were primed to different degrees, meaning that the difference in priming facilitation between priming and isolation could vary within the emotional categories. Figure 2 also shows the number of correct interpretations of the primes (multiplied by 10 for illustrative reasons), which correlated with RT difference for matching targets. This is further explained below in Experiment 3.

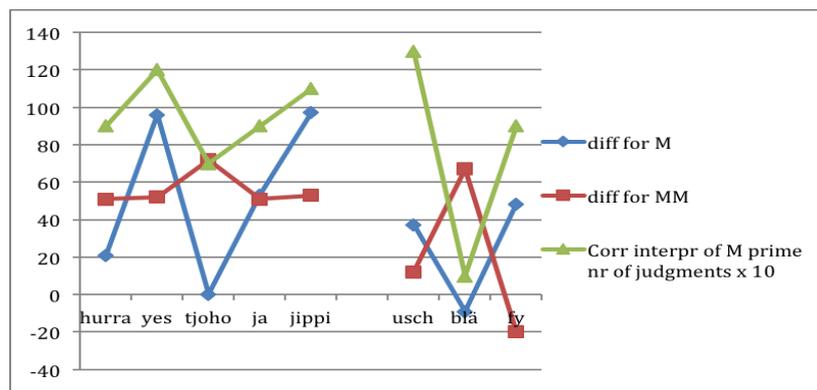


Figure 2. RT differences for matching and mismatching targets. Individual words were primed to different degrees. Correct interpretation of primes correlates with RT difference for matching prime (diff for M), but not with RT difference for mismatching prime (diff for MM).

RT difference between words in isolation and target words with matched prime was higher for five out of eight words, while RT difference between words in isolation and targets words with mismatched prime was lower, which can perhaps be explained by negative priming due to inhibition. The fact that mismatch gave rise to greater priming facilitation in three cases could be due to the fact that the corresponding primes were not sufficiently well produced in relation to the target.

As stated above in Section 2.1, there was no ‘internal priming’. Thus, the response times for all the targets were used in the analyses.

4.2. Experiments 1 and 2: Discussion

Experiments 1 and 2 revealed cross-modal priming from emotional prosody to written emotional interjections. The effect of the priming was significant for HAPPINESS interjections. The difference between primed and isolated targets was larger for HAPPINESS than for DISGUST interjections; that is, the primed HAPPINESS interjections had shorter RT. This is contrary to the results for non-human sounds in the study by Scherer and Larsen (2011), who found that negative auditory primes have a greater effect on visual word processing than do positive auditory stimuli. The result that HAPPINESS words were significantly primed, both in matched and mismatched condition, could point to HAPPINESS words being more easily primed by any emotional content. However, when we studied the individual words, another pattern emerged. Different words were primed to different degrees, depending on the accuracy of the prime prosody, and thus there was not an absolute difference of priming between HAPPINESS and DISGUST words.

5. Experiment 3: Method

In Experiment 3, the participants performed an interpretation test of non-words with emotional prosody.

5.1. Material

The stimuli consisted of 18 non-words with different emotional prosodies intended to express the emotions HAPPINESS, SADNESS, ANGER, SURPRISE, and DISGUST. After the interpretation test, two emotions were analysed: the results for the interpretations of the emotions HAPPINESS and DISGUST were extracted and thereafter compared with other variables. The stimuli tested were the same as in Table 1, column “Spoken non-word prime”.

5.2. Participants and procedure

The participants were 17 native speakers of Swedish (mean age: 24 years, 11 women). Participants listened over headphones to 18 non-words with different emotional prosody, and judged the emotional content of the expression in a forced-choice test, selecting between the interpretations HAPPINESS, SADNESS, ANGER, SURPRISE, and DISGUST.

5.3. Results

Figure 3 shows the number of expected interpretations of each non-word stimulus with emotional prosody.

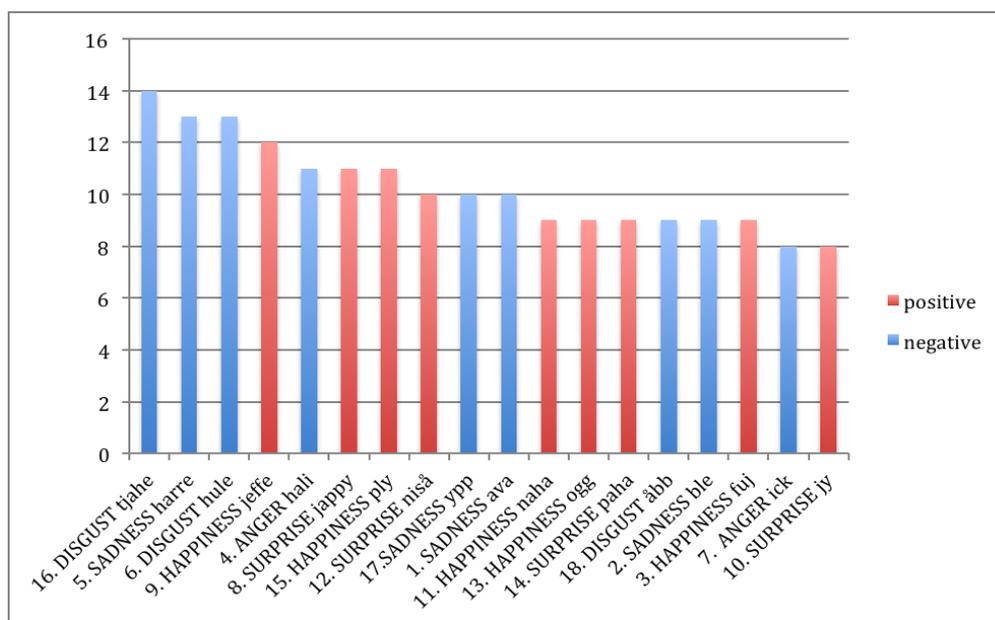


Figure 3. Interpretations of non-word stimuli with emotional prosody. The majority of the test-takers interpreted the emotions as the speaker intended. The figure shows the number of correct interpretations out of a possible 17.

As shown in Figure 3, most of the emotions were interpreted as intended by the speaker; for example, emotional expression ‘no. 16.’ was mainly interpreted as SADNESS (by 16 of 19 test-takers, i.e. by 84%), emotional expression ‘no. 5.’ was mainly interpreted as SADNESS (by 13 of 17 test-takers, i.e. by 76%), and emotional expression ‘no. 6.’ was mainly interpreted as DISGUST (by 13 of 18 test-takers, i.e. by 72%). The only exceptions were stimuli ‘no. 1. SAD’, which was interpreted as SURPRISED by 10 test-takers (but as SADNESS by 7) and ‘no. 14. SURPRISED’, which was interpreted as SADNESS by 9 test-takers (but as SURPRISE by 7). There was a slight tendency for negative emotions (DISGUST, ANGER, and SADNESS) to be better interpreted than positive ones (HAPPINESS and SURPRISE); in other words, there was slightly more consensus for the negative words than for the positive words. If we look at just the stimuli used in Experiments 1 and 2, namely prosodies for HAPPINESS and DISGUST, negative stimuli had 11.5 of 17 correct interpretations and positive had 10 out of 17 correct interpretations. The number of correct interpretations for primes expressing HAPPINESS and DISGUST is shown in Figure 2, in relation to the RTs for matched and mismatched prosody.

6. Study 4. Frequencies of target words

To discover the effect of word frequency, the frequencies for the eight target interjections were searched for in the Swedish language corpus KORP⁴, which includes a web-based tool for word search. KORP is a continuously growing corpus which uses written sources of written genres; at the time of our search it contained approximately 1 475 000 000 word tokens. There are several problems with frequency counts, in this case. First, KORP is based on written language, not spoken, while interjections are more common in spoken language. Another problem is that interjections are not particularly frequent in any corpus. However, a frequency count from KORP gives a suggestion of the (relative)

⁴ <http://spraakbanken.gu.se/korp>

frequencies of interjections. The most frequent interjection by far was *ja*, which had a frequency of only .05 % in the corpus.

The word frequencies extracted from KORP are shown in Table 2 and are discussed below in relation to Figures 4 and 5.

Table 2. Frequencies for the interjections in KORP

Interjection	Frequency in KORP
<i>Hurra</i>	1 648
<i>Yes</i>	5 609
<i>Tjoho</i>	568
<i>ja</i>	743 627
<i>jippi</i>	524
<i>usch</i>	9 856
<i>blä</i>	211
<i>fy</i>	12 173

7. Correlations of results

An analysis was performed of correlations between RT differences between isolation and priming, word frequency, and correct number of interpretations. Three variables were tested for correlations: (i) RT difference between priming and isolation, (ii) word frequency, and (iii) number of correct interpretations of prosody on non-words. There was a fairly high correlation between RT difference between matched priming and isolation, and word frequencies (.534). There was a significant correlation between RT difference between matched priming and isolation, and number of correct interpretations of prosody on non-words (.736). The correlations are shown in Table 3.

Table 3. Correlations between the four variables (A) RT difference between priming and isolation for matched prosody, (B) number of correct interpretations of prosody on non-words, (C) frequency, and (D) RT difference between priming and isolation for mismatched prosody for the targets *hurra*, *yes*, *tjoho*, *ja*, *usch*, *blä*, *fy*. M= matched, MM= mismatched.

B/A	C/A	B/D	C/D
Interpr/Diff	Frequency/Diff	Interpr/Diff	Frequency/Diff
M	M	MM	MM
0.736*	0.534	-0.458	-0.635

As shown in Table 3, there were positive correlations between correct interpretation of prosody and difference in priming time for matched words, and between word frequency and difference in priming time for matched words. The correlation between correct interpretation of prosody and difference in priming time for matching words was significant at the .005 level. Table 3 also shows that there were negative correlations between correct interpretation of prosody and difference in priming time for mismatched words, and between word frequency and difference in priming time for mismatched words, but these correlations were not significant.

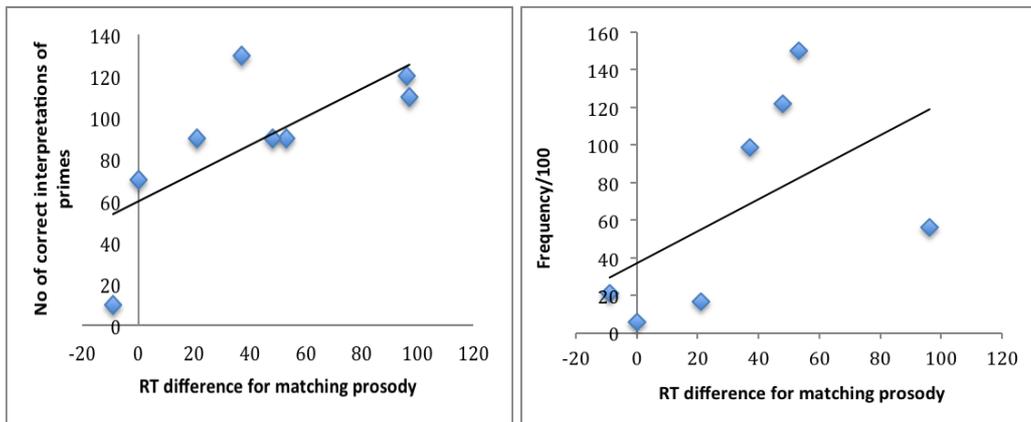


Figure 4. Correlations between RT difference between priming and isolation for matched prosody and number of correct interpretations (left) and RT difference between priming and isolation for matched prosody and frequency (right). Means are shown.

A Pearson correlation between (A) RT difference between priming and isolation for matched prosody and (B) number of correct interpretations of prosody on non-word primes produced a correlation of $r = .736$ which was significant at the .05 level. From this we can conclude that RTs decrease when the prime corresponds well with the emotional content in the target. There was no significant Pearson correlation between (A) RT difference between priming and isolation for matched prosody and (C) frequency.

On the other hand, there was a negative Pearson correlation of $r = -.458$ between (D) RT difference between priming and isolation for mismatched prosody and (B) number of correct interpretations of prosody on non-word primes. Although this was not significant, it points in the same direction as the result on facilitating matched prosody. Likewise, there was a non-significant negative correlation ($r = -.635$) between (D) RT difference between priming and isolation for mismatched prosody and (C) frequency.

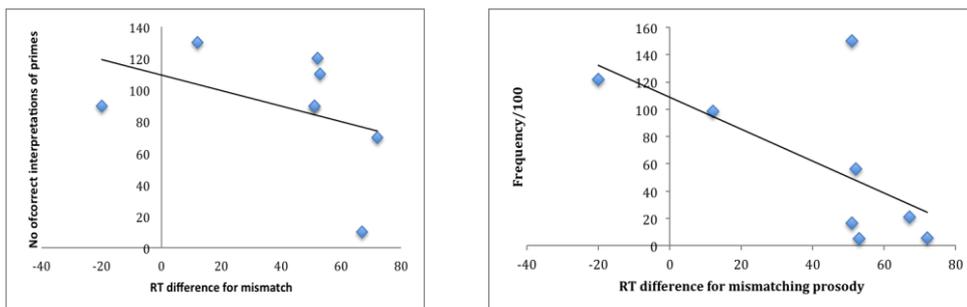


Figure 5. Correlations between RT difference between priming and isolation for mismatched prosody and number of correct interpretations (left) and RT difference between priming and isolation for mismatched prosody and frequency (right). Means are shown.

8. Conclusions and discussion

In some cases, the experiments revealed significant results, and in other cases there were tendencies. It is clear that more extensive studies need to be performed in this field, but some important conclusions can be drawn already. From Experiments 1 and 2 it can be

concluded that cross-modal priming takes place when the prime is a nonsense word with emotional prosody and the target is a written interjection with emotional content. There is a priming difference between positive emotion such as HAPPINESS and negative emotions such as DISGUST; as a group, positive emotions are primed more than negative ones. However, there were priming differences at the level of individual interjections, which could depend on the quality of the prime and the frequency of the interjections.

There was a correlation between priming facilitation and a measure of accuracy in interpretation of the primes. The relation can be interpreted as meaning that emotional prosodies that are more accurately produced in the prime stimuli activate the corresponding emotions and facilitate the recognition of the targets to a proportional degree.

There was also a tendency to a correlation between priming facilitation and word frequency. The effect can be interpreted as meaning that word frequency speeds up reaction times, as also seen in the reaction times for isolated stimuli, which is normally expected.

The findings have implications for the following theoretically relevant issues.

8.1. *Cross-modal priming and effects of different variables*

The experiments showed that there was an effect of cross-modal priming in the domain of study. How can this cross-modal priming of emotional prosody be explained? Matched emotional prosody of different valences (positive: HAPPINESS or negative: DISGUST) activates the corresponding emotion, and hence the recognition of the written emotional interjection is reacted to more quickly. When the emotional prosody is mismatched, there is inhibition. The best-interpreted emotional prosodies prime the corresponding emotional interjections to a higher degree. Thus, the priming is not a case of all or nothing; rather, it is gradual. There is also a frequency effect: more frequent emotions are more activated, and this is also seen in the higher degree of facilitation for these interjections.

The high proportion of correct interpretations of prosody in nonsense words shows that they are well produced for a particular emotional category. When they are well produced (and therefore well interpreted), they produce a greater priming effect on the written words in comparison with the experiment in which the written words were tested in isolation. The adequately produced emotional prosodies facilitate more rapid responses to the corresponding emotional interjections.

We thus interpret the results as follows. High frequency and correct production—and interpretation—gives greater facilitation in priming. Conversely, lower facilitation in priming, in the mismatched condition, does not correlate with high frequency or correct interpretation.

8.2. *Emotional prosodic morphemes*

The question now is whether the present results can be interpreted as semantic priming (emotion to emotion, irrespective of expression), morphological priming, or something else. In the case of morphological priming, it is possible to talk about ‘emotional prosodic morphemes’ (emoproemes), which are furthermore non-arbitrary in nature, just like phonaesthemes. As with the priming effects for different phonaesthemes (Abelin, 2012), there is not clearly a question of either/or, but of more or less.

8.3. Frequency

The results are in line with a usage-based theory of language (Bybee, 2010), considering the (near) correlation between degree of priming in targets and frequency of targets. In isolation, frequent words have faster reaction times than infrequent words, but there is also an increasing priming effect for more frequent word targets, which conspires with success of production. Since word frequency counts are laborious, an alternative possibility could be to use acceptability tests as a substitute for frequency counts in less-described languages. Other general problems with frequency counts have to do with global frequency versus frequency in context, and the relation between frequency and experienced familiarity.

8.4. Lexical gangs

The units of emotional prosodies and emotional meanings described as lexical gangs are sets of words with shared phonological and semantic properties that influence morphological productivity (Pierrehumbert, 2010). Phonaesthemes are a special case of lexical gangs, and emotional prosodic morphemes could constitute another. However, both phonaesthemes and emotional prosodic morphemes are not only explainable by frequency effects, but also have a non-arbitrary, indexical basis, which initiates their formation. They show gang behaviour, but with an indexical basis.

8.5. Non-arbitrariness

Where and how does non-arbitrariness enter into the discussion? In the present case, emotional prosody could be explained not only as learned (based on usage) but also as universal. The non-arbitrary relation between prosodic expression and meaning of emotional interjections is mainly *indexical*, in that the emotional prosodic expression is caused by a bodily or mental process. That is, there is a spatiotemporal, or even causal, relation between content (emotion) and expression, when considering the bodily aspect of emotion. By emphasizing the indexical relation between emotions and their expressions, we mean that the prosodic emotional expression emanates from a bodily reaction connected with emotion, not that the bodily reaction is the emotion. The effect of hypothalamic activation on the verbal expression of emotion is unclear,⁵ but phonetic research suggests that physiological changes in the articulatory musculature, resulting from hypothalamic activation and consequent arousal of the sympathetic nervous system, might affect the acoustic and perceptual characteristics of utterances made under emotional conditions (cf. Seddoh, 2002). Higher rate of airflow through the glottis associated with increased tension on the laryngeal musculature results in an increased rate of vocal fold vibration and therefore higher F0 values, and vice versa.

The biological basis of emotions is one of the roles of embodiment for emotional prosody. From another perspective, speakers express emotions indexically through interjections, and listeners experience these interjections with emotional prosody, beginning in childhood. Emotional prosody of interjections thus becomes coded in interjections through listeners' experiences. Embodiment and indexicality therefore go hand in hand, and emotional prosody can be considered part of language. Emotional prosody is also to some degree conventional, since it differs to some extent in different languages, as a consequence of learning. The high degree of priming of emotional

⁵ The hypothalamus is described as having a function of mediation of emotional responses, among other functions (Seddoh, 2002).

prosody found in our study is probably due to the non-arbitrary connection between expression and content.

The universal aspect of emotional prosody presents further evidence for this, since it is possible to interpret emotional prosody between languages to a certain extent. The question of universality of emotional expression and perception is not settled, although there seem to be more similarities than differences across languages (cf. e.g. Scherer, Clark-Polner, & Mortillaro, 2011). However, the fact that emotional prosody is universal to a great extent does not exclude it from being linguistic. Universality in syntax, phonology, or semantics does not exclude syntax, phonology, or semantics from the linguistic description, nor are adjectives or nouns expressing emotions excluded from description of language. We might need to dissolve the boundaries between ‘language’ and emotion, and consider the possibility that there are no clear-cut boundaries between ‘linguistic’ and emotional prosody. Non-arbitrariness is a crucial concept in linguistic theory, as the expression and content of emotional prosody are inseparably connected in interjections.

Acknowledgements

Warm thanks are extended to Isac Waern Kyrck who performed the lexical decision experiment of words in isolation (Experiment 2), and whose BA thesis (Waern Kyrck 2013) can be consulted for more details on this particular experiment.

References

- Abelin, Å. (1999). *Studies in Sound Symbolism*. Gothenburg Monographs in Linguistics 17. Göteborg.
- Abelin, Å. (2011). Perception of high-lexical and low-lexical interjections with matching and mismatching emotional prosody. *International Journal of Cognitive Linguistics*, 2(2), 127–142.
- Abelin, Å. (2012). Relative frequency and semantic relations as organizing principles for the psychological reality of phonaesthemes. Selected Articles from UK-CLA Meetings, Vol. 1 (pp. 128–145). UK-CLA.
- Abelin, Å. (2013). Durations and F0-variation of Swedish emotional interjections in relation to perception. *Nordic Prosody: Proceedings of the XIth Conference, Tartu 2012*, (pp. 49–58). Frankfurt am Main: Peter Lang.
- Alhner, F. & Zlatev, J. (2010). Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. *Sign System Studies*, 38(1/4): 298-348
- Becker, M. W. (2012). Negative emotional photographs are identified more slowly than positive photographs. *Attention, Perception and Psychophysics*, 74(6), 1241–1251.
- Bergen, B. K. (2004). The psychological reality of phonaesthemes. *Language*, 80(2), 290–311.
- Bybee, J. L. (2010) *Language, usage and cognition*. Vol. 98. Cambridge: Cambridge University Press.
- Johansson, N. & Zlatev, J. (this volume). Motivations for sound symbolism in spatial deixis: A typological study of 101 languages.
- Nobata, T., Hakoda, Y., & Ninose, Y. (2010). The functional field of view becomes narrower while viewing negative emotional stimuli. *Cognition and Emotion*, 24(5), 886–891.

- Nygaard, L. C., & Queen, J. S. (2008). Communicating emotion: Linking affective prosody and word meaning. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(4), 1017–1030.
- Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 325–347). Cambridge: Cambridge University Press.
- Pierrehumbert, J. B. (2010). The dynamic lexicon. In A. Chen, M. Huffman, & C. Fourgeron (Eds.), *Handbook of laboratory phonology* (pp. 173–183). Oxford: Oxford University Press.
- Scherer, L. D., & Larsen, R. J. (2011). Cross-modal evaluative priming: Emotional sounds influence the processing of emotion words. *Emotion*, *11*(1), 203–208.
- Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology*, *46*, 401–435.
- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research*, *14*(2), 228–233.
- Seddoh, S. A. (2002). How discrete or independent are ‘affective prosody’ and ‘linguistic prosody?’ *Aphasiology*, *16*(7), 683–692.
- Stankiewicz, E. (1964). Problems of emotive language. In T. A. Sebeok, A. S. Hayes, & M. C. Bateson (Eds.), *Approaches to semiotics* (pp. 239–264). London: Mouton.
- Waern Kyrck, I. (2013). The effect of valence on reaction times to Swedish emotional interjections (Bachelor’s thesis. Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg).
- Zsiga, E. C. (2014). *The sounds of language. An introduction to phonetics and phonology*. Oxford: Wiley-Blackwell.