

Das Trojansche Klassifikationsschema emotionaler Intonation und seine Anwendung auf Totenkopffaffenlaute¹⁾

U. Jürgens

Abstract. Trojan's classification system of human emotional intonations and its application on squirrel monkey calls. — According to Trojan (1975), human emotional intonations can be characterized by three parameters: loudness, pitch and pharyngolaryngeal tension (strain). Loudness is correlated with the sound pressure level, as measured in dB. Pitch is correlated with the fundamental frequency as well as the quotient of energy in the higher frequency band against that in the lower frequency band. Strain is represented by the amount of non-harmonic energy in phonation. Each parameter expresses a specific emotional dimension. Loudness indicates the appellative character of an utterance, that is, the degree to which the caller wants to change the behaviour of the partner. Pitch expresses the degree of control the caller has on the situation; low pitch (chest register) indicates good control, that is, self-assertiveness; high pitch (head register) indicates low control, that is, external events rather than internal plans determine the behaviour. Strained voice — in contrast to the relaxed one — indicates an unpleasurable state. A panic utterance, for instance, is characterized by great loudness, high pitch, great strain. Emotionally, it has a strong appellative character — representing, in fact, a cry for help. It expresses a loss of control over the situation and, furthermore, signals an aversive emotional state. Jubilating, on the other hand, is characterized by great loudness, high pitch, low strain. Again it has a strong appellative character: its aim is to invite others to join the jubilating person in its enthusiasm. The high pitch signals that the vocalizer feels overwhelmed by the event. The low strain indicates a non-aversive event. If situations in which squirrel monkeys (*Saimiri sciureus*) utter loud, high-pitched, strained calls or loud, high-pitched, unstrained calls or loud, low-pitched, unstrained calls etc. are compared with those in which humans produce emotional intonations with the same parameter constellation, a surprising correspondence in the emotional context is found. It is concluded that the vocal expression of emotion in humans has deep-reaching phylogenetic roots.

Key words. Emotional expression, prosody, vocalization, primates.

Aus kulturvergleichenden Untersuchungen ist bekannt, daß die Intonation, die beim stimmlichen Ausdruck bestimmter Emotionen verwendet wird, bei unterschiedlichen Kulturen sehr ähnlich ist. D. h. ein Amerikaner kann einen von einem Japaner wütend, ängstlich oder freudig gesprochenen Satz emotional richtig einordnen, auch wenn er kein Wort Japanisch kann und die Mimik des Sprechenden nicht sieht. Und das gleiche gilt umgekehrt (Beier & Zautra 1972). Diese transkulturelle Ähnlichkeit in der emotionalen Intonation spricht für eine genetische Grundlage solcher Äußerungen. Es stellt sich die Frage, wieweit diese genetischen Wurzeln in die Stammesgeschichte zurückreichen. D. h., wir wollen uns fragen, ob es Gemeinsamkeiten im stimmlichen emotionalen Ausdruck zwischen Affe und Mensch gibt.

¹⁾ Vortrag auf der 2. Arbeitstagung Säugetier-Bioakustik in Bonn.

Ich möchte mich zur Beantwortung dieser Frage eines von Trojan (1975) entwickelten Systems zur Klassifikation menschlicher emotionaler Intonationen bedienen. Trojan hat versucht, die emotionalen Intonationen durch 3 Gegensatzpaare zu charakterisieren: 1) Schonstimme — Kraftstimme; dieses Paar gibt an, wie laut oder leise gesprochen wird. 2) Bruststimme — Kopfstimme; hierin drückt sich aus, wie hell oder dunkel die Stimme klingt. 3) Rachenenge — Rachenweite; dies gibt an, wie gepreßt die Stimme klingt. Die 3 Gegensatzpaare bilden ein dreidimensionales Kontinuum, innerhalb dessen die verschiedenen Intonationen lokalisiert sind. Abb. 1 gibt tabellarisch acht Positionen innerhalb dieses Kontinuums wieder.

Jedem Gegensatzpaar ordnet Trojan eine emotionale Dimension zu. Durch das Schonstimme-Kraftstimme-Paar drückt sich aus, wie stark die Lautäußerung Appellcharakter hat, also eine aktive Reaktion vom Gegenüber erwartet wird. Durch das Gegensatzpaar Bruststimme-Kopfstimme drückt sich aus, wie selbstsicher bzw. Herr der Lage sich der Sprecher fühlt. Durch das Gegensatzpaar Rachenenge-Rachenweite, also den Grad des Gepreßtseins der Stimme, drückt sich aus, als wie unangenehm die augenblickliche Situation empfunden wird. Dementsprechend wird ein Angstschrei durch die Kombination Kraftstimme/Kopfstimme/Rachenenge repräsentiert, da er starken Appellcharakter hat (im Sinne eines Hilferufs), der Sprecher sich unsicher fühlt (Kopfstimme) und die Situation als unangenehm empfunden wird (gepreßte Stimme). Wütendes Schimpfen drückt sich ebenfalls in einer lauten Stimme aus, da es stark appellativen Charakter hat — versucht doch der Schimpfende Mitstreiter gegen die beschimpfte Person zu mobilisieren. Es wird mit gepreßter Stimme gesprochen, da die Situation als ärgerlich, also unangenehm, empfunden wird. Es wird im Gegensatz zur panischen Lautäußerung jedoch nicht in Kopf-, sondern Bruststimme gesprochen, weil das Schimpfen — als Ausdruck der Empörung — aus einer gewissen Selbstsicherheit heraus erfolgt, soll das Gegenüber doch eingeschüchert werden. Wenn wir dagegen von der Kombination Kraftstimme/Kopfstimme/gepreßte Stimme (wie sie für den Angstschrei charakteristisch ist) zur Kombination Kraftstimme/Kopfstimme/nicht-gepreßte Stimme übergehen, haben wir die jubelnde Stimme vor uns, die ebenfalls als Kraftstimme starken Appellcharakter hat — stellt sie doch eine Aufforderung dar, in die Begeisterung des Jubelnden mit einzustimmen. Die Kopfstimme drückt aus, daß der Jubelnde, weggerissen von der Begeisterung, wohl etwas die Kontrolle über sich und die Umwelt verloren hat, also mehr von der Situation bestimmt wird als sie selbst bestimmt. Die nicht-gepreßte Stimme drückt aus, daß der Jubelnde die Situation nicht als unangenehm empfindet. Entsprechendes gilt für die fünf übrigen Kombinationen in Abb.1.

Wenn wir diese Klassifikation auf Affen übertragen wollen, stellt sich zunächst das Problem, wie wir die 3 Gegensatzpaare Schonstimme-Kraftstimme, Kopfstimme-Bruststimme, Rachenenge-Rachenweite in objektive Meßgrößen übersetzen können. Obwohl dieses Problem noch nicht vollständig gelöst ist, können wir doch in erster Annäherung das Gegensatzpaar Schonstimme-Kraftstimme mit dem physikalischen Parameter Schalldruckpegel erfassen, das Gegensatzpaar Kopfstimme-Bruststimme durch die Grundfrequenz bzw. bei nicht-harmonischen Lauten durch das Verhältnis nieder- zu höherfrequenter Energie, und was das Gegensatzpaar gepreßte Stimme-nicht-gepreßte Stimme betrifft, durch das Verhältnis geräuschhafter zu harmonischer Energie. Wenn wir unter diesen Voraussetzungen das Trojansche Schema auf

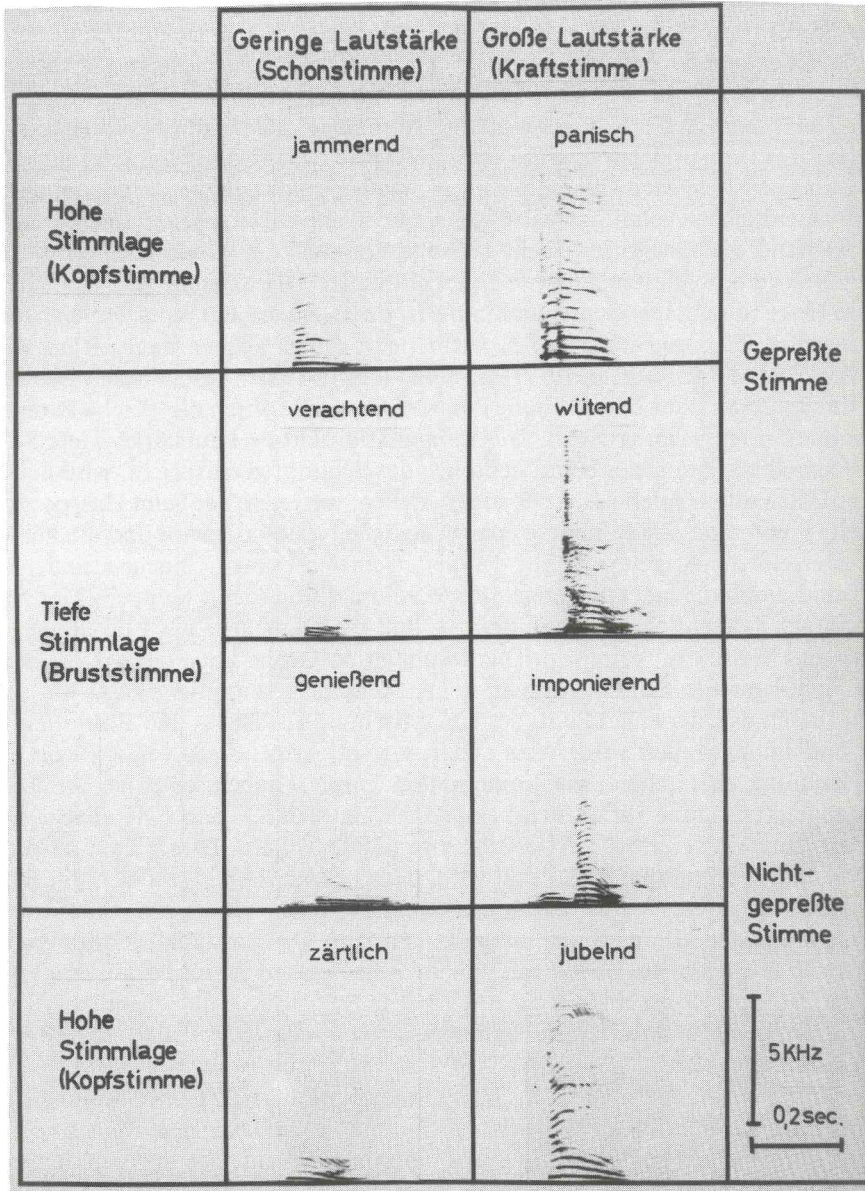
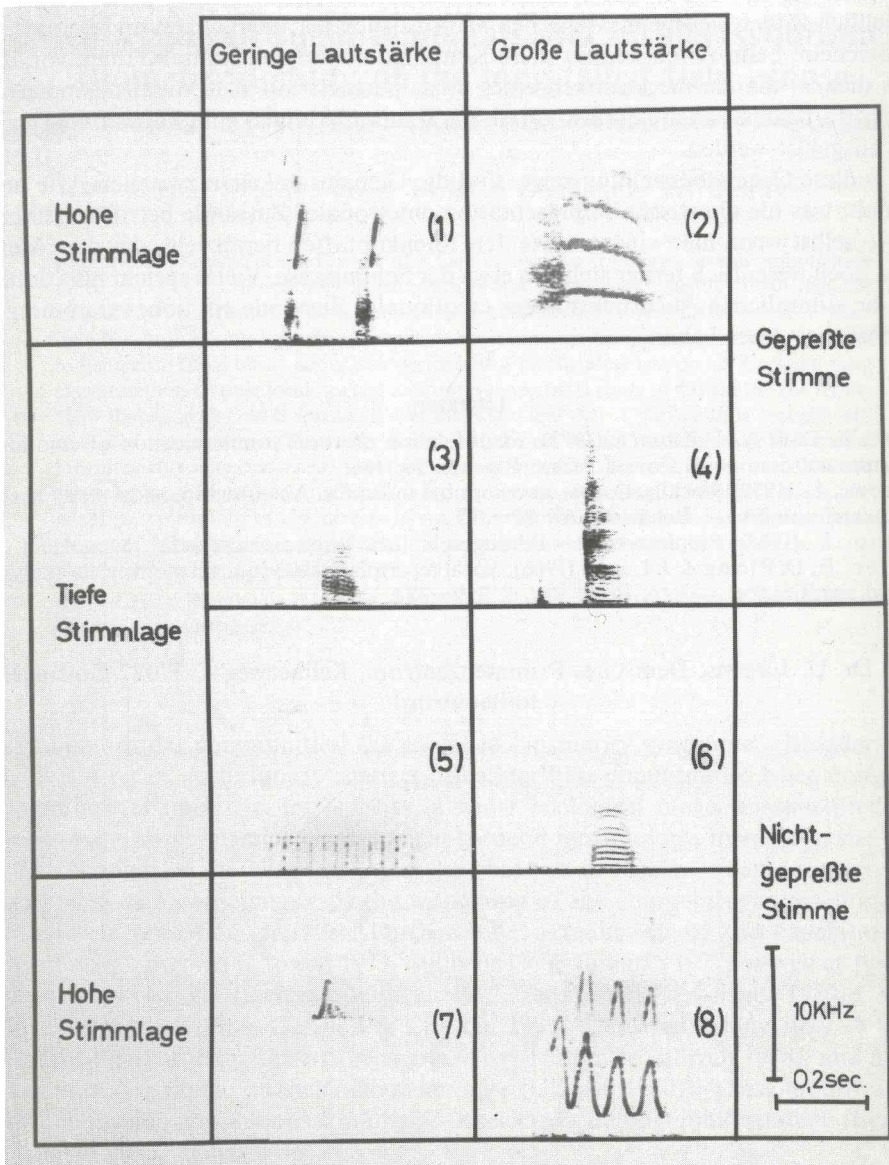


Abb. 1: Sonagramme des Wortes „Du“ in acht verschiedenen Intonationen vom gleichen Sprecher gesprochen.

die Laute von Totenkopffaffen (*Saimiri sciureus*), unseren Versuchstieren, anwenden, ergibt sich folgendes Bild:

In der linken Spalte von Abb. 2 sind Laute niedriger Intensität, genauer gesagt, unter 75 dB (bezogen auf $2 \times 10^{-5} \text{ N/m}^2$) dargestellt, in der rechten Laute über 75 dB; 75 dB entspricht etwa der Mitte des Gesamtdynamikbereichs von Totenkopffaffenlauten. In der Kategorie „Tiefe Stimmlage“ finden sich Laute, deren Hauptenergie unter 7 kHz liegt, in der Kategorie „Hohe Stimmlage“ solche mit Hauptenergie über 7 kHz; 7 kHz entspricht wiederum etwa der Mitte des Gesamtstimmumfangs dieser Tiere. In der Kategorie „Gepreßte Laute“ wurden solche Laute aufgenommen, die einen wesentlichen Anteil geräuschhafter, also nicht-harmonischer Energie aufweisen, während die Kategorie „Nicht-gepreßte Stimme“ nur Laute enthält, die keine deutlichen geräuschhaften Komponenten enthalten. Die Kombination Große Lautstärke/Hohe Stimmlage/Gepreßte Stimme ist durch einen Laut repräsentiert, der von unterlegenen Tieren geäußert wird, wenn sie von überlegenen angegriffen werden (Winter et al. 1966, Jürgens 1979). Er entspricht in seiner emotionalen Bedeutung also durchaus der panischen Stimme des Menschen, die durch die gleiche Parameterkombination repräsentiert wird. Die Kombination Große Lautstärke/Tiefe Stimmlage/Gepreßte Stimme, der beim Menschen das Schimpfen entspricht, wird beim Totenkopffaffen durch einen Laut, das sogen. Bellen, vertreten, der beim Hassen auf potentielle Freßfeinde, wie Raubtiere oder Schlangen, verwendet wird und die Funktion hat, einerseits die Artgenossen auf die Gefahr aufmerksam zu machen, andererseits den Feind durch die mächtige konzertierte Stimmaktion (denn wenn ein Tier zu belien anfängt, fallen die anderen mit ein — und die Lautstärke des Einzellautes liegt bei bis zu 110 dB!) zu vertreiben. Die Kombination Große Lautstärke/Tiefe Stimmlage/Nicht-gepreßte Stimme drückt beim Menschen Imponiergehabe aus. Beim Affen finden wir hier einen Laut, der eine Mischung aus Droh- und Protestlaut darstellt und dann geäußert wird, wenn es darum geht, Artgenossen von einer bestimmten Handlung abzuhalten. Die Kombination Große Lautstärke/Hohe Stimmlage/Nichtgepreßte Stimme findet sich beim Menschen beim Jubeln. Die gleiche Merkmalskombination zeigt beim Affen das sogenannte Trillern. Dieser Laut wird geäußert, wenn eine Futterquelle entdeckt wird, oder wenn Gruppengenossen, die längere Zeit getrennt waren, wieder zusammenkommen, oder wenn nach einer längeren Periode bedeckten Himmels die Sonne durchbricht. Trillern signalisiert also wie Jubeln freudige Ereignisse. Es hat mit Jubeln außerdem gemeinsam, daß es stark ansteckend ist, d. h. wenn ein Tier zu trillern anfängt, ist die Wahrscheinlichkeit groß, daß auch andere Tiere zu trillern beginnen — selbst wenn das Trillern nur vom Tonband dargeboten wird.

Die Kombination Geringe Lautstärke/Hohe Stimmlage/Gepreßte Stimme finden wir bei einer jammernden, weinerlichen Stimme. Beim Totenkopffaffen finden wir mit der entsprechenden Merkmalskombination einen Laut, der nur von Jungtieren geäußert wird, wenn diese von der Mutter gesäugt werden wollen, die Mutter sich jedoch entzieht. Geringe Lautstärke, tiefe Stimmlage und gepreßte Stimme finden sich beim Menschen beim Ausdruck der Verachtung, beim Affen bei Lauten, die ein leichtes Sich-Unwohlfühlen bei geringer Aktionsbereitschaft ausdrücken, also ein Unwohlsein, das weder zu deutlichen Flucht- noch zu Aggressionshandlungen Anlaß gibt, sondern das Unangenehme als gegeben hinnimmt. Geringe Lautstärke, tiefe

Abb. 2: Sonagramme von acht verschiedenen Lauttypen des Totenkopffaffen (*Saimiri sciureus*).

Stimmelage und nicht-gepreßte Stimme findet sich beim Menschen bei Lauten des behaglichen oder wollüstigen Genießens, beim Affen in Form sogenannter Schnurr-laute. Diese treten z. B. auf beim Paarungsvorspiel, aber auch zwischen Mutter und Kind beim Säugen. Geringe Lautstärke, hohe Stimmelage, nicht-gepreßte Stimme schließlich finden sich beim Menschen als Ausdruck der Zärtlichkeit und beim Einschmeicheln; beim Affen kommt diese Kombination bei den Kontaktlauten vor, die dazu dienen, die Aufmerksamkeit eines Sozialpartners auf den Vokalisierenden zu lenken. In beiden Fällen soll also durch die Lautäußerungen eine Zuwendungsreaktion ausgelöst werden.

Wie diese Gegenüberstellung zeigt, sind die Gemeinsamkeiten zwischen Affe und Mensch, was die akustische Repräsentation emotionaler Zustände betrifft, erstaunlich — selbst wenn man eine Art wie den Totenkopffaffen heranzieht, der dem Menschen doch wesentlich ferner steht als etwa der Schimpanse. Vieles spricht also dafür, daß die stimmlichen Ausdrucksmuster emotionaler Zustände ein hohes stammesgeschichtliches Alter haben.

Literatur

- Beier, E. G. & A. J. Zautra (1972): Identification of vocal communication of emotions across cultures. — *J. Consult. Clin. Psychol.* 39: 166.
- Jürgens, U. (1979): Vocalization as an emotional indicator. A neuroethological study in the squirrel monkey. — *Behaviour* 69: 88–117.
- Trojan, E. (1975): Biophonetik. — Bibliograph. Inst. Wissenschaftsverlag, Mannheim.
- Winter, P., D. Ploog & J. Latta (1966): Vocal repertoire of the squirrel monkey, its analysis and significance. — *Exp. Brain Res.* 1: 359–384.

Prof. Dr. U. Jürgens, Deutsches Primatenzentrum, Kellnerweg 4, 37077 Göttingen.