

Jochen Raecke

**„SYNCHRONISIEREN ODER UNTERTITELN?“ - DAS IST HIER DIE
FRAGE. WESHALB ES ZUR LINGUISTIK DER SPRACHE EINE
LINGUISTIK DES SPRECHENS HINZU BRAUCHT**

1. Wie die Frage und ihre Begründung in und auf die Welt kamen

Was im Titel dieses Beitrags als Frage im Stile des berühmten Hamletmonologs formuliert ist, also die Frage, „ob's edler im Gemüt" sei, dass einem Zuschauer, der von allem, was in einem Film gesprochen wird, nichts versteht, das für ihn Unverständliche in "seiner" Sprache ans Ohr gebracht wird oder dass man einem solchen Zuschauer das für ihn Unverständliche in "seiner" Sprache unter die Augen bringt, diese Frage also, die – anders und genauer gefasst – darauf hinausläuft, ob man Menschen, denen man gern einen Film zeigen möchte, obwohl sie alles Sprachliche darin gar nicht verstehen würden, diesen Film besser dadurch verständlich zu machen versucht, dass man allen Schauspielern, die darin etwas zu sagen haben, zunächst einmal eine andere Stimme gibt und sie dann mit dieser anderen Stimme auf der Leinwand einen anderen Text sprechen lässt, dabei aber einen Text, von dem die gewünschten Zuschauer glauben können, es sei ein Text ihrer Sprache, oder ob man solch "unverständigen" Menschen derartige Filme nicht besser in der Weise verständlich zu machen versucht, dass man sie zwar hören lässt, was die Schauspieler in der ihnen unverständlichen Sprache von sich geben, ihnen am unteren Bildrand aber gleichzeitig in ihrer Sprache zu lesen gibt, was die Schauspieler damit sagen wollen, diese erste Frage im Titel des vorliegenden Beitrags also tritt im Jahre 2010 bereits in ihr 83. Lebensjahr.

Was sodann nach dieser Frage im Titel zwar mit einem Fragepronomen beginnt, aber gar keine Frage darstellt, sondern eine Erklärung dafür, weshalb ein solcher Beitrag ausgerechnet in einem Sammelband linguistischer Beiträge erscheint, kann als Desiderat auf ein noch höheres Alter zurückblicken, denn jenes Werk, das die Linguistik der Sprache als die eigentliche und allein richtige Sprachwissenschaft begründete, der *Cours de linguistique générale* von Ferdinand de Saussure erschien bekanntlich 1916 in Genf. Mit der Unterscheidung von *langue* und *parole* und der These, dass nur die *langue* der wirkliche Gegenstand der Linguistik sei, weil sie dem Sprechen zugrunde liege, wurde das Sprechen im Prinzip aus der Linguistik verbannt. Indem das Sprechen einfach als

Realisierung einer *langue*, genauer des Systems einer Sprache angesehen wurde und damit in Vergessenheit geriet, was seinerzeit Wilhelm von Humboldt gesehen hatte, dass nämlich die Sprache real nur als Sprechen existiert und ihrer Natur nach ein reines Konstrukt ist, resp. etwas, was seiner ursprünglichen und wesenhaften Sinnlichkeit vollständig entkleidet ist. Wie ich an anderer Stelle versucht habe, aufzuzeigen und zu begründen, hat sich die Sprachwissenschaft im Laufe der Zeit immer mehr zu einer reinen Schriftsprachenwissenschaft entwickelt und dabei aus dem Blick gelassen, dass Menschen ihre Fähigkeit, sich sprachlich zu äußern, ausschließlich am Beispiel anderer Menschen entwickeln, die gerade mit ihnen sprechen. Jegliche Form von Schriftsprachlichkeit ist damit sekundär. Das heißt nicht, dass man sich damit nicht beschäftigen sollte, man sollte sich dessen aber klar bewusst sein (Raecke 1998). Ist man dieses, so läuft die Frage "Synchronisieren oder Untertiteln" auf eine andere und durchaus linguistische Frage hinaus, nämlich die, ob denn das „Verständlichmachen“ unverständlicher sprachlicher Äußerungen in schriftlicher Form und das Verständlichmachen sprachlicher Äußerungen in mündlicher Form nicht so verschiedene Dinge sind, dass man zu ihrer Beschreibung auch verschiedene theoretische Ansätze braucht, vor allem dann, wenn der "Unverständige" sieht, wie eine ihm unverständliche sprachliche Äußerung von einem Sprecher oder einer Sprecherin produziert wird. Übersetzt man nämlich eine solche Äußerung in eine schriftliche Form, dann erfährt man in dieser nur, was von der betreffenden Sprecherin oder dem betreffenden Sprecher gerade gesagt wird und alles sinnlich wahrnehmbare an dieser Äußerung geht in dieser Übersetzung vollständig verloren, übersetzt man eine solche Äußerung aber so, dass ein "Unverständiger" den Eindruck gewinnen soll, die gerade zu hörende Äußerung werde in der ihm verständlichen Sprache gerade von den entsprechenden Personen produziert, dann geht es zwar auch darum, was gerade gesagt wird, aber nicht allein darum, sondern um erheblich mehr. Es muss das gerade zu Hörende auch zu dem passen, was im Gesicht, am Körper und an den Bewegungen der jeweilsprechenden Personen abzulesen ist. Und da wissen wir aus unserer Kultur, dass jemand, der mit dem Kopf nickt, eher ein Ja als ein Nein von sich gibt, und dass jemand, der mit dem Kopf schüttelt, eher Ablehnung als Zustimmung signalisiert. Gesichtsausdruck, Haltung und Bewegung sind also mit sprachlichen Inhalten korreliert und vor allem synchronisiert, und also muss die Übersetzung eigentlich in das hineingesprochen werden, was gerade an der sprechenden Person zu sehen ist. Womit klar werden sollte, dass das Synchronisieren und das Untertiteln es mit verschiedenen Dingen zu tun haben, einmal allein mit dem, was die Sprache allein leistet, nämlich einen rein geistigen Inhalt mitzuteilen, und das andere Mal mit dem, was das Sprechen als konkreten, beobachtbaren körperlichen Vorgang von der Sprache als Transportmittel für nicht beobachtbare abstrakte Inhalte unterscheidet. Das heißt zusammengefasst und auf den

Punkt gebracht, dass der vorliegende Beitrag über die Diskussion der Frage, ob Synchronisieren oder Untertiteln “edler im Gemüt” sei, zugleich und in Sonderheit ein abermaliges Plädoyer für eine Linguistik des Sprechens liefern will, die ich schon einmal im Rahmen einer Konstanzer Linguistik (Raecke 1998) als ebenso sinnvoll wie notwendig zu begründen versucht habe.

2. Wenn der Ton mehr als die Musik macht

Um noch einmal mit der Frage im Titel und ihrem Alter anzufangen, so ist sie rein logisch betrachtet auf den Tag so alt, wie der Tonfilm heute alt ist, und der kam bekanntlich – wenigstens nach heute geläufiger Ansicht – mit dem Streifen *The Jazz Singer* 1927 auf die Welt. ‘Natürlich’ in den Vereinigten Staaten von Amerika. Auch wenn das Wenige, das in diesem ersten Tonfilm tatsächlich mit Ton *gesprochen* wurde – mit Ton wurde in diesem Film im Wesentlichen *gesungen*; was im Drehbuch als ‘Gesprochenes’ ausgezeichnet war, erschien, wie bis dahin üblich, als Zwischentitel – gar nicht zum Synchronisieren gedacht war. Denn wenn man Wikipedia (Wikipedia) einfach einmal glauben darf, dann waren Monologe und Dialoge darin nur improvisiert, weil Warner Bros. gar nicht *mehr* wollten, als einen Film produzieren, in dem *Musik* und *Gesang* synchronisiert wurden. Entsprechend war auch kein zu synchronisierendes Dialogmanuskript notwendig und was im realisierten Film dann doch mit Ton *gesprochen* wurde, war improvisiert. Es gab zwei solcher Sprachsequenzen, von denen die erste Al Jolsons, des Hauptdarstellers, allgemeine Lieblingssequenz war, die er auch auf der Bühne regelmäßig von sich gab, nämlich: *Wait a minute! Wait a minute! You ain't heard nothin' yet. You wanna hear Toot-toot-tootsie? Al right.* Die zweite Sequenz war um einiges länger, aber im gleichen Stil und trug im Prinzip genauso wenig zum Verständnis des ‘eigentlichen’ Films bei wie die erste.

Machte dementsprechend bis 1927 der Ton die Musik *zum* Film, änderte sich das mit dem *Jazz Singer* grundlegend, denn jetzt machte die Musik den Ton *im* Film. Und von da aus war es dann nur noch ein Trippelschritt, bis der Ton mehr als nur die Musik *im* Film machte, er machte schließlich den ganzen Film – wenn man denn die Integration des Sprachlichen in die bewegten Bilder so in Worte kleiden darf. Ganz in dieser Weise aber sah den Umstand, dass die Sprache jetzt in *gesprochener* bzw. zu *hörender* Form Teil der Bilder selbst wurde, wenigstens der zu seiner Zeit und noch heute berühmte Filmkritiker Siegfried Kracauer, als er am 4.12. 1930 die Premiere des Films *Im Westen nichts Neues* in Berlin erlebte. Der inzwischen als ‘Klassiker’ gehandelte Film von 1930 war der Versuch, den weltbekannten gleichnamigen Roman von E. M. Remarque für ein englischsprachiges Kinopublikum einzurichten, denn er entstand in den Vereinigten Staaten. Für den deutschen Sprachbereich wurde er lippensynchronisiert und flimmerte so am besagten Tage am besagten Ort über die Leinwand. Was

Kracauer von dieser herab in die Augen fiel und aus den Lautsprechern zu Ohren kam, muss ihm wie ein Sturmangriff auf seinen guten Filmgeschmack vorgekommen sein. Denn er feuerte in der fälligen Besprechung dieses Ereignisses auf der ersten Seite seiner *Frankfurter Zeitung* scharf zurück, indem er schrieb: "Der Versuch, amerikanische Schauspieler für deutsche auszugeben, ist ein Unding." und gab zu bedenken, dass "der tönende Film die Internationalität des stummen" nur dadurch bewahren könne, dass "man entweder das Schwergewicht von den Dialogen zurück auf die Bilder oder auch auf die Geräusche verlegt oder jeden Film von vorn herein in allen Hauptsprachen" dreht (Kracauer 1930).

Damit erkannte Kracauer gleichsam schon bei der Geburt dieser Art von Film in ebenso bewundernswerter wie unerbittlicher Schärfe, dass mit ihr aus dem Film insgesamt etwas anderes werden würde, als er bis dahin gewesen war. Im Unterschied zu vielen anderen, die darin, dass die Bilder jetzt nicht nur das Laufen, sondern auch noch das Sprechen gelernt hatten, eine technische Meisterleistung sahen, von dem der Film nur großes Heil – vor allem natürlich auch großes finanzielles Heil – erwarten konnte, sah Kracauer darin eher Unheil heraufziehen. Auf jeden Fall richtig sah er damit, dass die Bilder im tönenden Film – was genauer zu lesen ist als: mit der zu *hörenden* Sprache – ihren Alleinvertretungsanspruch resp. ihre Selbstgenügsamkeit aufgaben. Und das hieß nichts anderes, als dass der Film da, wo er wirklich *nur* Film oder *reiner* Film war, nämlich *zwischen* den Zwischentiteln, seine Universalität (Kracauer sagt wörtlich allerdings nur 'Internationalität') verlor. Denn Bilder, die keine Sprache in sich zeigen, sind theoretisch oder potentiell von jedem Menschen auf dieser Welt zu verstehen. Mit der Sprache als integralem Bestandteil jedoch bekommen sie *nolens volens* Lokalkolorit, Regionalkolorit oder – und das unausweichlich – Nationalkolorit. Sie werden entsprechend lokalisiert, regionalisiert und vor allem nationalisiert. Das bedeutet, dass ihr 'richtiges' oder 'wirkliches' Verständnis an die Kenntnis entsprechender Lokalitäten, Regionen oder Nationen gebunden ist. Und über alles dieses hinaus werden die Bilder durch den Ton auch noch individualisiert. Denn mit dem 'tönenden Film', in welchem Sprecher so sprechen, dass man es nicht mehr nur, wie im Stummfilm, *sehen*, sondern auch zeitgleich *hören* kann, wird der (Spiel-)Film als Genre zugleich sprachen- wie sprecher-spezifisch.

Da es die gesprochene Sprache nur in einer zahlenmäßig gar nicht zu bestimmenden Vielfalt gibt, die – ganz genau betrachtet – geradezu unendlich viel größer ist, als die ja auch schon nicht geringe Vielfalt der geschriebenen Sprachen, wird das, worüber man sich auf der einen – und zwar der kleineren – Seite nur freuen kann, zu etwas, über das man sich auf der anderen – und zwar der wesentlich größeren – Seite nur ärgern kann. Freude bringt es nur denen, die 'ihre' Schauspieler jetzt genau so sprechen hören können, wie sie es in dieser Rolle

realiter getan haben, Ärger bringt es all denen, die die Schauspieler zwar sprechen hören können, davon aber eigentlich gar nichts haben, weil sie sie nicht verstehen. Und da die Zahl derer, die einen Schauspieler dann, wenn er seine Muttersprache spricht, wirklich, vollständig und richtig verstehen können, wesentlich geringer ist als die Zahl derer, die einen Film mit ihm sehen wollen oder sollen, ist der Ärger, den sich der seinerzeit ja noch recht junge Film mit dem Ton eingehandelt hat, gleichsam vorprogrammiert.

Gemindert werden kann er nur durch das, was zur Überwindung von Verständigungsschwierigkeiten zwischen Sprechern verschiedener Sprachen schon von altersher eingesetzt wurde, nämlich das Übersetzen. Aber was soll bei gesprochener Sprache eigentlich genau übersetzt werden? Übersetzt man nur, was gesagt wird, geht verloren, wie es gesagt wird. Und da nun grundsätzlich nur übersetzt werden kann, was *gesagt* wird, nicht aber, wie *gesprochen* wird, ist auch die Übersetzungswissenschaft relativ bald mit ihrem Latein am Ende. Das *Sprechen* von Personen kann schließlich nur imitiert werden. Und wenn in anderen Sprachen *gesprochen* wird, so muss es zwar zunächst übersetzt werden, muss dann aber auch wieder imitiert werden. Also sollte die Übersetzung so sein, dass sie eine möglichst gute Imitation ermöglicht. Ob 'wörtlich' oder 'äquivalent' ist hier gerade nicht die Frage.

Was einem als gesprochene Sprache an die Ohren kommt, ist folglich nur *so* in eine andere Sprache zu übertragen, dass eine Unzahl von Kompromissen geschlossen, eingegangen oder in Kauf genommen wird. Weshalb man es auch so gut wie niemandem, der das Original kennt, recht machen kann. Der eine findet diesen Kompromiss einfach faul, jener findet ihn immerhin erträglich, der andere kann jenen Kompromiss gerade noch billigen, diesen aber nicht mehr. Irgend etwas gibt es immer auszusetzen. Auch wenn man scherzhaft zu bedenken gäbe: das Original gibt es nur einmal, hörte man gereizt zurück: es gibt aber bessere und schlechtere Kopien.

Schlussendlich könnte man so weit gehen, in der Integration des Sprachlichen – vorbereitet in seiner geschriebenen Form, endgültig in seiner gesprochenen Form – den Sündenfall des Films, will sagen, den Anlass für die Vertreibung aus dem Paradies der bewegten Bilder zu sehen. Denn mit dem Synchronisieren – und dem Untertiteln kaum weniger – müssen nun viele Beteiligte tatsächlich und geradezu wörtlich ihr Brot im Schweiß ihres Angesichts essen. Weil sie sehr hart arbeiten mussten, um es zu verdienen.

Kracauers Gedanken verdienen es noch heute, als Ausgangspunkt für ausgedehnte geistige Spaziergänge auf diesem weiten Feld genommen zu werden. Er zeigte sich in diesen Gedanken als mehr denn bloß arroganter Filmkritiker, sich elitär gebärdender Pessimist oder konservativer Verteidiger des Hergebrachten, wie er verschiedentlich gesehen wurde, er zeigte sich vielmehr als jemand, der einfach tiefer dachte, als viele Zeitgenossen und Filmfreaks es tun, denen es im

Kino in lautlicher und sprachlicher Hinsicht nicht naturgetreu genug zugehen kann. Die von Kracauer prophezeite Lokalisierung, Regionalisierung und Nationalisierung des ursprünglich universellen Mediums Film ist schließlich Wirklichkeit geworden, wird gerade beim Synchronisieren und beim Untertiteln immer wieder schmerzlich bewusst, ja es macht diese überhaupt erst notwendig. Und sie sind es, die denjenigen, die auf diesem Gebiet auf der Jagd nach finanziellem Erfolg sind, ständig unanständig dicke Knüppel zwischen die Beine werfen und sie dadurch – sehr salopp formuliert, aber am besten ins Bild passend – kräftig auf die Schnauze fallen lassen.

Nun will ich in diesem Monolog allerdings nicht von solchen praktischen Schwierigkeiten schwadronieren. Die sind auf der einen Seite zwar wirklich unterhaltsam mit Beispielen zu illustrieren und immer wieder lehrreich – wie z. B., dass man in einem Untertitel lesen konnte: *Speichert unsere Seelen* und erst langsam realisierte, dass im Original 'Save Our Souls' gesagt wurde –, denn man kann sich dann immer wieder theatralisch kopfschüttelnd und ungefährdet besserwisserisch fragen, wie jemand mit solchen Untertiteln versorgt einen Film verstehen soll, hier will ich viel mehr dem 'eentlichen' Grundproblem, das Kracauer in der Tiefenstruktur seiner Äußerung angesprochen hat, nachgehen und aufzeigen, dass es im mehr als komplexen, nämlich durchaus problematischen Verhältnis von Wort und Bild liegt. Darüber wird bei diesem Thema – wenn es denn überhaupt seriös angesprochen wird – eher selten nachgedacht. Nur liegen da aus meiner Sicht die Wurzeln des Problems von Synchronisieren und Untertiteln. Einerseits des Problems, das sie beide zusammen darstellen und sie zugleich vereint, andererseits ihrer jeweiligen 'persönlichen' Probleme. Kommt man an seine Wurzeln, so sollte sich das Problem ganz von selbst in einen festeren geistigen Griff begeben, als es der ist, in dem man es heutzutage – wenigstens nach meinem bisherigen Eindruck – hat..

3. Wenn zusammenwächst, was nicht zusammengehört

Als Kracauers argumentative Grundlage nehme ich ein klares Wissen darüber an, dass das Bild von Natur aus und von seinem Ursprung her keine Sprache kennt. Alles Sprachliche ist ihm wesensfremd, weil es sozusagen die Welt anders anschaut und entsprechend anders repräsentiert als die Sprache. Die Höhlenmalerei zeigt das eindrücklich. Man kann nicht *irgendeine* Beute, also *irgend etwas* zu Essen an die Wand malen, man kann nur entweder einen Hirsch, ein Bison oder sonst etwas konkret Seiendes und damit sichtbar in Erscheinung Tretendes an die Wand malen. Was man sieht, ist immer nur ein bestimmter von allen möglichen Hirschen oder Bisons. Der Betrachter kann sie – womöglich erst heute – als sichtbare Verkörperungen *des* Hirsches oder *des* Bisons *deuten, sehen* tut er – auch heute noch – immer nur diesen oder jenen *so* oder *so* gemalten.

Was mit den Augen erfasst wird, ist seiner Natur nach immer konkret und individuell. Sprachliche Ausdrücke sind dieses gerade nicht, sie sind immer abstrakt. Selbst Personennamen. Denn man kann leicht sagen: "Da kommt schon wieder so ein Peter." Das unterscheidet sprachliche lexikalische Einheiten von der Welt, wie sie ist, denn mit der Sprache wird die Welt erfasst bzw. strukturiert, sie wird mit ihr aber nicht abgebildet. Am einsichtigsten macht das die für viele Eltern befremdliche Tatsache, dass Kinder eine bestimmte Zeit lang alle Frauen und alle Männer, die sie wahrnehmen, als *Mama* und *Papa* bezeichnen. In der Welt, wie sie ist, begegnet der Mensch ausschließlich Individuellem. Jeder Stein für sich ist anders als jeder ihm noch so ähnliche andere Stein und jeder Fluss für sich ist anders als jeder andere Fluss. Eigen-, Orts- oder Personennamen stellen zwar einen Versuch dar, diesem Faktum der grundsätzlichen Individualität in der Sprache Rechnung zu tragen. Gelingen kann er aber nicht. Weil sprachliche Zeichen zur Bezeichnung von etwas Seiendem symbolische Zeichen sind und keine ikonischen. Der Sinn liegt in dem, was die Sprache leisten soll: Wir sollen uns mit ihr in der Welt zurechtfinden. Und da können Einzelheiten nur stören. Bekäme jedes Einzelne seinen ihm von seiner realen Einzelheit her ja zustehenden eigenen Namen, müssten wir Millionen von Namen im Kopf haben. Aber wo sollte der Sinn liegen, jedem Hirsch und jedem Hasen im Wald einen Namen zu geben, nur weil es davon keine zwei gleichen gibt? Das, was wir die Bedeutung von Wörtern nennen, sind mehr oder weniger starke Abstraktionen von dem, was die mit einem jeweiligen einzigen Wort zusammengefassten einzelnen Seienden voneinander unterscheidet. Als Beweis reicht der Blick darauf, was alles mit jeweils dem gleichen Wort *Baum*, *Vogel* oder *Hund* zu einer bestimmten Art von Erscheinung zusammengefasst wird.

Das heißt sehr einfach, dass Bild und Sprache ursprünglich nicht zusammengehören, weil Bilder nichts anderes tun, als stumm etwas zu *zeigen*, meist angefertigt mit der Idee, dass Betrachtende darin etwas wiedererkennen können oder sollen. Vom Menschen mit der freien Hand oder auch mit der Kamera geschaffene Bilder werden im Gehirn der Betrachtenden abgeglichen mit dem, was in diesem Gehirn als bereits gesehen abgespeichert ist und zum Zwecke der Verständigung über sie durch ein von allen anderen Wörtern verschiedenes Wort unterschieden wird. Wenn Bilder denen, die sie betrachten, jedoch mehr als nur etwas *zeigen*, nämlich etwas *sagen* sollen, dann braucht es dazu mehr als dieses einen Wortes, das nur bezeichnet, was man sieht, es braucht dazu wenigstens zweier Wörter und damit das, was wir die Sprache nennen, mit einer Grammatik, die wissen lässt, welches dieser zwei Wörter dasjenige ist, das das in Rede stehende Etwas nur nennt, und welches von ihnen dasjenige ist, das dem Genannten etwas zuschreibt, d.h. darüber etwas aussagen soll. Bilder haben keine Gedanken, sie bringen uns höchstens auf solche. Und die können wir – wie

Humboldt schon erkannte – nur äußern, weil wir die Sprache haben (Humboldt 1836/1960).

Dabei hat das Bild mit der Anschauung zu tun, das Wort mit der Vorstellung. Und auch wenn Vorstellungen sich auf Anschauungen gründen, sind sie dennoch wesensverschieden. Maler tragen dem meist dadurch Rechnung, dass sie die Deutung des von ihnen Dargestellten durch einen Titel einschränken. Aber der befindet sich grundsätzlich außerhalb des Bildes. Darüber oder darunter. Wilhelm Busch z.B. hat dieser ursprünglichen Wesensfremdheit der Sprache im Bild Rechnung getragen – wie seine weniger bekannt gewordenen Zeichnerkollegen auch –, indem er alles das, was seine Bilder kommentierte und verständlich machte, in Versform unter sie setzte. Die Sprache hatte für ihn im Bild nichts zu suchen und blieb deshalb außen vor. Wenn Ikonen oder andere Bilder religiösen Inhalts Schriftzeichen *innerhalb* des Dargestellten aufweisen, so scheint das zwar auf den ersten Blick ein Widerspruch zu der These zu sein, dass sich das Bild der Sprache verweigert, aber dieser Widerspruch löst sich schon auf den zweiten Blick, wenn man realisiert, dass alles Sprachliche auf Ikonen oder anderen 'Heiligenbildern' lediglich *Namen* sind, und also nicht aus "normalen" Wörtern besteht. Das erklärt auch, weshalb diese Namenwörter – wenigstens auf orthodoxen Ikonen – meist in einer besonderen Schreibweise erscheinen, nämlich abgekürzt unter sog. Titeln. Das deshalb, weil über diese Schreibung der PN (P[ersonenname]n) signalisiert oder gar dokumentiert wird, dass es sich im vorliegenden Falle eben gar nicht um *Bilder*, sondern um *Abbilder* der oder des Dargestellten handelt. Über ihren Namen nämlich werden die Dargestellten als einmalig, individuell und zugleich konkret indiziert. Was Betrachtende sehen sollen, ist nicht ein Bild, das man sich entgegen dem alttestamentarischen Bilderverbot von dem jeweils Dargestellten gemacht hätte – das könnte dann nämlich 'abstrakt' im Sinne der Visualisierung einer reinen Vorstellung sein –, sondern ein *Abbild*, das in ungebrochener Tradition auf jemanden zurückgeht, der den oder die Abgebildeten in seiner fleischlichen Gestalt so gesehen hat. Sei es real oder in einer von Gott gesandten Vision.

Was nun den Stummfilm mit seinen Unter- bzw. Zwischentiteln angeht, so war er Wilhelm Busch in dieser Auffassung von der Sprache noch gefolgt. Nur der Comic brach mit ihr. Ganz zu Beginn der Serie *The Yellow Kid* (1897) trägt der Held, also *Yellow Kid*, alles, was er (oder 'es') denken oder sagen soll, in Worte gefasst auf seinem Hemd vor sich her. Oder auch mit sich herum, womit dann zugleich im Prinzip auch die Sprechblase in die Welt gekommen ist (Fuchs, Reitberger 1983, 21). Was wiederum zunächst in den USA praktiziert wurde, entwickelte sich leicht erkenntlich zum Zeichentrickfilm, der ja – wie bei den ersten Mickey-Mouse-Filmen nun wirklich in die Augen springt – letztlich gar nichts anderes ist als ein Comic mit bewegten Bildern und einer Tonspur an-

stelle von Sprechblasen. Der Tonfilm geht also auf den Comic zurück, und deshalb ist er wohl auch in Amerika entstanden.

Dennoch bleibt es für den Film ganz allgemein dabei: Im Anfang war das Bild, und das Bild war im Film, und der Film war das Bild. Es ist das Bild, das den Film zum Film macht und es sind nicht die heute z.T. atemberaubende Handlung, nicht der z.T. brillante Dialog, nicht die z.T. geniale Filmmusik (etwa von Enrico Morricone). Denn wenigstens Handlung und Dialog machen von sich aus noch keinen Film, sondern zunächst nur Theaterstücke. Und die Musik dazu würde höchstens eine Oper machen. Auch wenn diese drei selber natürlich schon mit dem (Bühnen-)Bild und damit unter Einbeziehung visueller Elemente arbeiten. Und so gibt es tatsächlich Filme – vor allem dann, wenn das Drehbuch von Menschen stammt, die sonst viel eher von der Handlung oder von der Sprache her mitreißende Stücke für die Bühne schreiben –, die viel weniger Film sind als auf die Leinwand versetztes Theater. Da sind die Bilder gerade gut genug, die Dialoge personell und räumlich zu situieren und zu illustrieren, nicht aber, für sich zu sprechen und durch das Gesprochene nur interpretiert zu werden.

Gleichwohl scheint in heutiger Zeit, wie Kracauer es befürchtet hatte, der eigentliche Bildcharakter des Films im Bewusstsein stark in den Hintergrund getreten zu sein, wenn nicht gerade pure *Action* ins Bild gesetzt wird, bei der ja in aller Regel höchstens hysterisches Geschrei, aber kein 'wohlgeformter Satz' mehr an die Ohren dringt. Der weitgehende Verzicht auf wirkungsmächtige Bilder, wie sie z.B. Eisenstein in seinem *Panzerkreuzer Potemkin* exemplarisch auf die Leinwand gebracht hatte, könnte sich daraus erklären, dass das Verhältnis von Bild und Film gar kein so ungebrochenes ist, wie oben suggeriert wurde, sondern sogar ein sehr gebrochenes, indem der Film ja erst entsteht, als die Bilder *laufen* lernen. Damit werden nämlich die Grenzen, die dem Bild gesetzt sind, überwunden, das Einzelbild verliert seinen Rahmen und damit zugleich seine phänomenologische Gebundenheit an den Moment. Lessing hatte bekanntlich in seinem 'Laokoon' (Lessing 1766/1971) für die bildende Kunst insgesamt festgestellt, dass es in ihr gelte, den entscheidenden Moment festzuhalten, während die Sprache Entwicklung und damit Handlung darstelle. Indem der Film jedoch Bilderfolgen von der Art bietet, dass sie dem Auge sie Betrachtender vorgaukeln, die Dinge auf der Leinwand bewegten sich, nähert sich der Film – so paradox das nach dem bisher Ausgeführten klingen mag – der Sprache an. Natürlich bleibt es beim entscheidenden Unterschied der sprachlichen Zeichen, aber mit den bewegten oder laufenden Bildern tut sich die Möglichkeit auf, das zu tun, was man *erzählen* nennt. Erzählen heißt wörtlich und seinem Wesen nach immer: 'erstens das, zweitens das, drittens das', jedoch mit der Maßgabe, dass das Erste zu Ende ist, ehe das Zweite eintritt, und das Zweite zu Ende ist, wenn das Dritte beginnt. Beim Erzählen geht es folglich um das *Nacheinander*

und so etwas gibt es auf *dem* Bild nicht. Dort gibt es nur ein Nebeneinander, also die Gleichzeitigkeit im strengsten Sinne. Die Serialität von Bildern im Film ändert alles. Und da nun erstens ein zeitliches Nacheinander unter bestimmten Bedingungen auch Geschehen oder Bewegung genannt wird, und zweitens menschliche Bewegungen normalerweise Handlungen, d.h. intendierte Zustandsveränderungen von etwas, sind, wird für den Film schon nach ganz kurzer Zeit die Möglichkeit entdeckt, genauso Geschichten zu erzählen, wie es über die Sprache möglich ist. Genauer gesagt: etwas so darzustellen, dass die Betrachtenden das Gesehene als eine Geschichte wiedergeben würden.

Wenn nunmehr mit den bewegten Bildern etwas möglich geworden ist, was man mit der Sprache auch kann, wozu braucht man dann aber noch die Sprache? Die Erklärung liegt in dem, was man entweder gar nicht oder jedenfalls nicht gut zeigen kann. In Faßbinders *Effi Briest* Verfilmung gibt es eine Reihe von Szenen, die geradezu paradigmatisch deutlich machen: Das Bild kommt einem Menschen nicht näher als bis zu den Augen. Was hinter den Augen liegt oder anfängt, ist nicht abzubilden. Womit wir wieder bei der prinzipiellen Konkretheit des Bildes sind. Wenn Menschen handeln und damit die Voraussetzung für eine Geschichte schaffen, haben sie für die einzelnen Handlungen in der Regel das, was man Motive nennt. Solche liegen aber ausschließlich im Kopf oder im Charakter eines Menschen vergraben, über Bilder sind sie nicht zu zeigen, sondern höchstens sinnfällig zu machen. Dies jedoch so zu tun, das es eindeutig ist, gehört zum Schwierigsten, was ein Regisseur sich zumuten kann. Deshalb ist es leicht verständlich, dass sehr bald nach der Erfindung des Films – als seine technischen Möglichkeiten noch viel weniger entwickelt waren als heute – nach solchen Möglichkeiten gesucht wurde, die es erlaubten, die Sprache zu den Bildern hinzukommen könnte. So allerdings, dass das Sprachliche auf die Bilder bezogen wird und allein zu dem Zweck, dass die prinzipielle Offenheit in der Deutung der Bilder eingegrenzt wird. Weil so die Folge der Bilder verständlich wird.

Wer eine Geschichte allein mit Bildern erzählen will, muss sich klarmachen, dass er ausschließlich die Augen ansprechen kann. Für diese, also die Augen kommt eine Geschichte nur dadurch zustande, dass etwas, das als identisch mit einem auf dem vorigen Bild Dargestellten erkannt oder gedeutet wird, eine sichtbare Veränderung durchgemacht hat. Im Film – mit bewegten Bildern also – kann solche Veränderung nun in der Weise dargestellt werden, dass Betrachtende sie als ununterbrochen empfinden. Die Identität des Dargestellten ist folglich kein Problem mehr: wenn nichts unterbrochen wird, muss das zu Sehende immer das Gleiche sein. Und so kann man jetzt einen Zug in einen Bahnhof einfahren lassen, ihn dort anhalten lassen, man kann eine in Schwarz gekleidete Frau aussteigen lassen, die von einem ebenfalls in Schwarz gekleideten Mann in den Arm genommen wird und wenn jetzt nach einem Schnitt eine Trauerfeier

gezeigt wird, auf der die beiden in ihrer schwarzen Kleidung sozusagen unverändert anwesend sind, dann hat man damit eine sprachlose Bildergeschichte erzählt. Allerdings müsste die Trauerfeier gezeigt werden, damit die Zugfahrt der Dame und ihre Begrüßung auf dem Bahnhof für Betrachtende wirklich so verstanden wird, dass hier eine entfernt wohnende Verwandte oder Bekannte zur Beerdigung eines Anverwandten oder Bekannten angereist ist. Würde die Trauerfeier nicht gezeigt und sie wäre tatsächlich der Anlass für den Besuch der Dame in Schwarz, dann müsste von ihr gesprochen werden. Und das heißt jetzt ganz schnell und formelhaft gesagt: Alles, was im Film nicht – wörtlich – offensichtlich ist, auch wenn er im Unterschied zum einfachen Bild die Bewegung zur Verfügung hat, muss Zuschauenden über die Verwendung von Sprache als die einzig richtige Deutung deklariert werden.

Genau dieses ist mithin die Funktion der ursprünglichen Unter- oder Zwischentitel. Und wenn man es genau betrachtet, liegt hier eigentlich auch ein möglicher Ansatz, zu einer rationalen Bestimmung dessen zu kommen, was Untertitel im Tonfilm ganz generell leisten sollen. Sie sollen genau das leisten, was ursprünglich die Sprache im Film zu leisten hatte, nämlich auf keinen Fall Selbstzweck zu sein, indem Personen aus ihren Worten heraus charakterisiert werden sollen, sondern das Gezeigte oder zu Sehende verständlich werden zu lassen, indem die prinzipiell offenen Deutungsmöglichkeiten der Bilder eingegrenzt werden. Der Ausgangspunkt des Films ist es, die Welt, die Menschen und ihre Schicksale vor Augen zu führen, will sagen, aus dem heraus zu deuten oder verständlich werden zu lassen, was man von ihnen sehen kann. Mit der Sprache konnte man das längst vor der Erfindung des Films, aber alles, was über die Ohren verarbeitet wird, lässt eben nur subjektive Bilder im Kopf entstehen, keine objektiven für die Augen. Und insofern sollte sich die Sprache im Film eigentlich grundsätzlich als lediglich komplementär begreifen, d.h. nur da eingreifen wollen, wo das Verständnis von Bildern seine natürlichen Grenzen hat. Respektive die Leistung von Bildern ihre Grenzen hat und die Sprache die einzige Möglichkeit ist, eine Botschaft eindeutig zu vermitteln.

Das ist nun auch Kracauers früher Appell gewesen, dass man “das Schwergewicht von den Dialogen zurück auf die Bilder” legen sollte, also formelhaft und kämpferisch: zurück zum Bild, nieder mit der Sprache. Wenn aber Dialoge im Sinne des Gesagten sozusagen unvermeidlich sind, dann sollte man sich in Untertiteln eben darauf beschränken, das in ihnen zu zeigen – denn mit Buchstaben kann man ohnehin nicht mehr als zeigen –, was der Unverständige für das Verständnis gerade zu Sehenden braucht, und das ist das, was im strengen Sinne des Wort *gesagt* wird. Was gesprochen wird, ist individuell, lokal, regional oder national, und das lässt sich allenfalls beim Synchronisieren wiedergeben, wenn man es denn will.

4. Wenn die Sprache im Film auch noch eine persönliche Note bekommt

Der fundamentale Unterschied zwischen dem Tonfilm und dem Stummfilm mit seinen Zwischentiteln liegt nun, ausführlicher gesagt, darin, dass Sprachliches im letzteren visuell, im ersteren aber akustisch vermittelt wird. Indem damit die Sprache im Tonfilm aus dem Medium des Visuellen in das Medium des Akustischen wechselt, bekommt jede und jeder, die oder der einen solchen 'tönenden' Film anschaut, erheblich mehr Informationen über die Darsteller geliefert, als man sie ihr oder ihm im Stummfilm mit den geschriebenen Texten je hätte bieten können. Denn zum eigentlichen Inhalt, also dem Gesagten, gesellen sich in der akustisch wahrnehmbaren Version auch alle diejenigen Informationen, die beim Sprechen allein über das Ohr wahrgenommen werden können. Zu betonen ist, dass es dabei um ein Phänomen der *Wahrnehmung* geht. Denn *vermittelt* werden können solche Informationen auch über geschriebene Sprache. Nur macht es einen erheblichen Unterschied, ob zu *lesen* ist, er spreche deutlich den Südstaatlerakzent, oder ob das zu *hören* ist. Im einen Falle muss man sich den Südstaatlerakzent vorstellen, im anderen Falle nimmt man ihn wahr, selbst dann, wenn man gar nicht weiß, dass es der Südstaatlerakzent ist, den man hört. Man nimmt aber immer wahr, dass der Sprecher 'anders' spricht, als diejenigen Schauspieler im gleichen Film, die den Standard sprechen. Dagegen gilt: wer diesen Akzent nicht kennt, kann ihn sich auch nicht vorstellen.

Das gibt Gelegenheit, die von Kracauer realiter zwar nur indirekt angesprochene, aber durch die Natur der Sache eben gegebene *Lokalisierung*, *Regionalisierung* und *Nationalisierung* des Films durch den Ton etwas genauer zu illustrieren, indem die Augen noch einmal auf vorstehend zitierte Sequenz aus dem *Jazz Singer* gelenkt werden. Wer genügend *American English* kann, der weiß, dass sie zunächst einmal ein geradezu klassisches Beispiel für das sog. *Black English* ist, das in keiner anderen Sprache der Welt adäquat wiedergegeben werden kann, weil es eine solche 'schwarze' Varietät in anderen Sprachen nicht gibt. Denn *Black English* (*BE*) ist ein soziales Phänomen oder auch eines der Hautfarbe, weshalb es dort, wo es zu Hause ist, keine regionale oder lokale Erscheinung ist, wie es z.B. Dialekte im Deutschen wären. Wer versuchte *BE* durch Bairisch, Schwäbisch oder Berlinerisch wiederzugeben, erntete im besten Falle einen Lacherfolg, es deswegen aber in sog. Gossensprache wiederzugeben, wäre genauso 'daneben'. Darüber hinaus gibt es aber von diesem *BE* nun auch wieder eine unüberschaubare Menge lokaler Varianten, und wenigstens ein südstaatliches *BE* ist deutlich ein anderes als ein nordstaatliches. Zugleich hat aber auch New Orleans sein besonderes *BE*. Welches *BE* Al Jolson genau imitierte, muss hier nicht diskutiert werden, dass es allerdings wirklich nur Imitation war, ist jedoch ebenso sicher wie von Bedeutung. Al Jolson selber war nämlich gar nicht 'black', sondern sehr 'white'. Schließlich war er der Sohn eines jüdischen

Arztes aus Litauen. Nur pflegte er die heute meist als rassistisch abgelehnte, damals aber sehr beliebte Masche des *blackface*, bei der man sich als Weißer das Gesicht schwarz färbte und als Neger gab, d.h. sie imitierte oder viel häufiger karikierte. Und das natürlich auch sprachlich. Für die jeweils Anwesenden bei den Bühnenshows war es allerdings von großer Bedeutung, ob es um Südstaatler *BE* oder Nordstaatler *BE* ging, weil man sich mit dem nordstaatlerischen *BE* eher mit den Schwarzen solidarisierte, während man sich mit dem südstaatlerischen von ihnen distanzierte. Entsprechend wichtig wäre dies natürlich auch im Film gewesen, denn der Tonfilm macht aus denen, die gerade zuschauen, ja nicht nur Augenzeugen, sondern zugleich auch Ohrenzeugen, und damit in einem ganz anderen Sinne Anwesende als der Stummfilm. Bei dem Versuch, *BE* in der Schrift anzudeuten, gehen über das Ohr leicht wahrzunehmende Unterschiede jedoch aus praktischen Gründen verloren. Denn wer nur zuschaut und nicht zuhört, erfährt in der Schrift allein, dass Al Jolson *BE* spricht, er erfährt aber nicht, wie er es spricht. Die Schrift kommt nämlich über Stilisierung nicht hinaus. Es sei denn, man transkribiere wissenschaftlich.

Wer noch tiefer blickt, muss jedoch zu diesen lokalen, regionalen und nationalen Komponenten, die durch die Sprache der Darsteller in einen Film notwendigerweise hineinkommen, noch hinzuerkennen, dass bekanntlich jeder Mensch eine Stimme hat, die nur ihm allein eignet, die ihn unverwechselbar macht und die ihn genauso als Individuum ausweist, wie seine Fingerabdrücke. Von daher kommt mit dem Ton zu allen sozialen, regionalen, nationalen Zügen auch noch das Phänomen der stimmlichen *Individualität* hinzu. Ein Phänomen, das dem Film bis dahin völlig fremd gewesen war, weil Druckbuchstaben keine individuellen Züge tragen. Die jeweilige Handschrift der Sprecher wäre zwar so individuell gewesen, wie ihre Stimme und ihre Art zu sprechen, sie wäre in der Regel aber *zu* individuell gewesen, als dass man sie so schnell hätte lesen können, wie der Film in seinem Drang nach Vorwärts das verlangt. Was in der Schrift also nur theoretisch hätte sinnvoll sein können, war mit dem Tonfilm ganz automatisch praktisch gegeben. Jedes Sprechen hat eine persönliche Note. Und die bekommt es bzw. behält es jetzt auch im Film.

5. Wenn Verstand (wieder einmal nur) Leiden schafft ('Gore ot uma')

Zugleich wird aus Kracauers früher Kritik aber auch noch ein Weiteres deutlich, das nämlich, dass für fast jeden, der die Sprache eines sog. Originalfilms versteht, sowohl das Synchronisieren als auch das Untertiteln ein Kreuz sein muss. Ob synchronisiert oder untitled wird, ihm einerlei, er muss Trauer tragen um das Original und kann nur leiden unter dem, was ihm im einen wie im anderen Falle in einem so bearbeiteten Film als Mangel, Ungeschicklichkeit oder gar Fehler in die Augen oder in die Ohren sticht. Wieder einmal schafft Verstand nur

Leiden, wie Griboedov in seinem *Gore ot uma* ja mit einem teuflischen Lachen sinnfällig gemacht hat. Man kommt also, wenn man es bildlich noch einmal etwas anders fassen will, bei der praktischen Lösung der Frage, ob "Synchronisieren oder Untertiteln", zugleich nicht weiter als vom Regen in die Traufe. Hat man aber nur die Wahl zwischen zwei Übeln, tritt literarisch gebildeten Menschen unweigerlich Hamlet im Kopf auf und fragt, welches von beiden Übeln man als das größere ansehen solle oder wolle, oder entsprechend auch als das kleinere.

Ein rational, logisch und pragmatisch denkender Zeitgenosse würde hier ohne Zögern sagen: Wenn man es denn wüsste, so wäre das kleinere Übel die bessere Lösung und das größere Übel gehörte ganz einfach abgeschafft. Ein von Lessing kritisch aufgeklärter Zeitgenosse sagte dagegen etwas anderes: Wenn sich doch in dieser Frage, ob es edler im Gemüt sei, zu synchronisieren oder zu untertiteln, die Geister so stark scheiden, dass etwa in Deutschland 78 Prozent aller einschlägig Befragten für das Synchronisieren sind und nur 4 Prozent für das Untertiteln, in Holland dagegen 82 Prozent aller einschlägig Befragten für das Untertiteln sind und sich nur 12 Prozent für Lippensynchronisation erwärmen können (Pruys 1997, 15), möchte es dann nicht so sein, dass beide Seiten jeweils so viele gute Gründe für ihre Entscheidung auf ihrer Seite haben, dass es eine Lösung für diese Frage gar nicht geben kann? Wenn der 'Herr des Ringes' also wie in Lessings Ringparabel gar kein 'Herr des Ringes', sondern ein 'Herr der Ringe' wäre? Von denen in diesem Falle allerdings jeder Ring auch seine Macken hat, dabei aber keiner mehr und keiner weniger als der andere? Wäre es da nicht weiser, nur der Frage nachzugehen, worin die Macken denn bestehen und wie sie da notwendigerweise hineinkommen? Denn dies würde in jener grandios paradoxen Weise, die uns Lessing in seiner Ringparabel als der eigentlichen Weisheit letzter Schluss vorgeführt hat, den Weg zu einer Lösung öffnen, die zugleich eine Lösung und doch keine ist. Die 'Lösung' bestünde nämlich darin, darauf zu verzichten, nach einer solchen zu suchen. Denn die Weisheit des Nathan liegt ja in dem nach der Vernunft einzig möglichen Appell, aufzuhören, in einer Frage Krieg zu führen, auf die es keine Antwort im Sinne des berühmten Kierkegaard'schen *Entweder – Oder* gibt, sondern nur eine Antwort vom Typus *sowohl als auch*. Im hier verhandelten Falle hätte die Antwort etwa den Wortlaut: "Beides, das Untertiteln wie das Synchronisieren, ist zwar schlechter als das Original, aber ohne das eine oder auch ohne das andere könnten ganz viele Menschen mit dem Original gar nichts anfangen." Und deswegen erscheint es – kritisch aufgeklärt und dem dialektischen Prinzip der Einsicht in die Notwendigkeit folgend – pragmatisch allein sinnvoll, nachzudenken oder besser: nachzuschauen, worin zunächst einmal überhaupt die Notwendigkeit, das eine oder das andere zu tun, gründet, was man beim einen wie beim anderen sowohl verliert als auch gewinnt, und wie man, dann Schiller folgend: "Der Not gehorchend, nicht dem ei-

genen Antrieb folgend“ dieses notwendiger Weise verlorengelassene möglichst gering hält oder möglichst gut erträglich macht. So, dass der Bedürftige bekommt, was sein Bedürfnis stillt. Eingedenk der Tatsache, dass Medizin die Gesundheit nicht ersetzen kann. Sie lindert immer nur die Symptome für das Fehlen der Gesundheit.

Wenn im Übrigen gerade gesagt wurde, man komme bei der praktischen Lösung der Frage ‘Synchronisieren oder Untertiteln’ nicht weiter als vom Regen in die Traufe, dann klingt das, so direkt gesagt, ziemlich defätistisch, und man könnte zu Kracauers Schluss kommen, man sollte doch auf beides verzichten und alle Filme wenigstens in allen Hauptsprachen drehen. Nur hat Kracauer in diesem Punkt wirklich nur ins Blaue hineingedacht, denn was er vorschlug oder gar forderte, hieße ja: Alle Menschen, die Filme sehen wollen, müssten entweder wenigstens eine Hauptsprachen so gut lernen bzw. beherrschen, dass sie darin alle Feinheiten der Dialoge mitbekommen. Das Thema würde sich damit in der Tat erübrigen und kein Mensch mehr den Wunsch verspüren, sich in einem wissenschaftlichen Beitrag öffentlich darüber Gedanken zu machen. Nur verrät der Blick in die graue Wirklichkeit, dass Kracauers Idee die Produzenten von Filmen viel zu teuer käme. Man stelle sich vor, die Hunderte von Millionen, die z.B. *Lord of the Rings* in der Herstellung gekostet hat, sollten auch für das Französische, Spanische, Portugiesische, Russische und noch einige Hauptsprachen ausgegeben werden. Wie sollten die bei dem dann ja verkleinerten Markt wieder hereinkommen? Und wer sollte sie entsprechend überhaupt hineinstecken wollen? Sollen außerdem kleinere Länder dann nur noch Filme in fremden Sprachen drehen? Soll man sich wirklich vorstellen, in Italien würde ein Film mit italienischen Schauspielern gedreht, die alle englisch sprechen? Wer wollte das verstehen, und wer wollte sich das wirklich 90 Minuten lang anhören? Verzichtete man aber auf das Eine wie auf das Andere, so hieße das, dass ein sehr großer Teil von Filmen allein in ihren Entstehungsländern sinnvoll gezeigt werden könnte. Serbische Filme z.B. könnten praktisch nur von Serben gesehen werden, wenn wir die inzwischen auch nicht mehr so vielen Kroaten und Bosniaken nicht zählen, die nicht verlernt haben wollen, Serbisch im Kino zu verstehen. Und was wäre mit kroatischen Filmen, mit bosnischen und montenegrinischen?

6. Wenn eine Frage über achtzig und keiner ein bisschen weise wird

Weil es damit auf jeden Fall sinnvoller erscheint – um es ihm Rahmen eines anderen Mediums ins Bild zu setzen –, ein Buch mit Druckfehlern in die Hand zu bekommen und es mit solchen lesen zu können, als dass man es gar nicht in die Hand bekommt, weil die Hersteller Angst vor Druckfehlern haben, wird vor allem in Deutschland seit der Geburt des Tonfilms fleißig synchronisiert und in anderen Ländern, z.T. weil es billiger ist, z.T. aber auch aus anderen Gründen,

seit dieser Zeit fleißig unternimmt. Und das bedeutet weiterhin, wenn man es noch einmal etwas anders betrachtet, dass die Frage 'Synchronisieren oder Untertiteln' in unseren Tagen bereits in ihr Greisenalter eintritt.

Gegen mögliche Erwartungen bedeutet dieses Alter aber nicht, dass die Frage deshalb schon sehr viel in einer für sie zuständigen Wissenschaft erlebt hätte, erlebt in dem Sinne, dass ihr Wissenschaftler ein lebhaft zu nennendes Interesse entgegengebracht oder ihr gar in schriftlicher Form den Hof gemacht hätten. Auf solche Menschen, gleich welcher Couleur wirkte sie bislang alles andere als attraktiv, was man daran sieht, dass sie sich mit ihr bis heute nur selten ernsthaft einlassen wollten und sie mehrheitlich ziemlich weit links liegen ließen. Von theoretisch folgenreichen Affären mit ihr ist entsprechend wenig zu hören, zu lesen oder zu berichten. Und wenn Pruys (1997, 22-42) das mit natürlich ganz anderen Worten auf genau 21 Seiten vor nunmehr auch schon wieder 12 Jahren ausführlich referiert, zu Recht kritisiert und darüber sehr verständlich rät, so hat das wenig oder gar nichts geändert. Auch nach Pruys' Buch sah sich kaum jemand veranlasst, dieser Frage, die in der alltäglichen Praxis für so viel Aufregung, Betriebsamkeit und Diskussionen sowohl in den Filmverleihen, Synchronstudios als auch vor allem bei den Filmkritikern sorgt (Pruys 1997, 16), entspannt und theoretisch fundiert tief in die Augen zu sehen oder an den Ohrflüppchen zu spielen. Was sich in der vergleichsweise spärlichen Literatur findet, soll hier auch aus Platzgründen nicht noch einmal referiert werden, der vorliegende Beitrag greift es da auf, wo es hinpasst. Auch wenn man den Eindruck gewinnen kann, dass Wissbegierigen im Wesentlichen nur Dinge bewusst gemacht werden, die sie leicht hätten selber erkennen können, wenn sie die Augen resp. die Ohren etwas weiter aufgemacht hätten. Und deshalb kann man mit einem nicht ganz wörtlich von Kurt Jürgens übernommenen Wort sagen: die Diskussion, ob's edler im Gemüt, zu synchronisieren oder zu unternimmt, ist inzwischen mehr als achtzig Jahre alt, aber die Wissenschaft ist in Bezug auf sie immer noch "kein bisschen weise".

7. Wenn jedoch Charme und Reiz mit dem Alter nicht verloren gehen

Wer sich nun aber trotz ihres fortgeschrittenen Alters auf eine Affäre mit dieser Frage einlässt, kann überrascht sein, wie viel Charme sie eigentlich noch immer hat und welchen Reiz sie auf einen ausüben kann. Wenn man sie denn dazu bringen kann, einem etwas zu sagen, und sie dann auch reden lässt. Dann vermag man sich lange nicht mehr aus ihrem Zauberkreis zu lösen und das auch dann, wenn man langsam merkt, dass sie einen am Ende wirklich viel eher in den Regen stellt. Oder unter die Traufe?

Was diese Frage, um es noch einmal ein Spur salopper zu sagen, alles zu bieten hat und wie oder womit sie einem die Sinne und den Verstand betören kann,

das lässt sich im Rahmen eines Aufsatzes natürlich nur sehr unvollständig beschreiben. Allerdings meinte ja Goethe schon in seiner unerschöpflichen Weisheit, erst in der Beschränkung zeige sich der Meister, und hat damit genauso recht wie die Volksweisheiten: 'in der Kürze liegt die Würze' bzw. 'manchmal ist weniger mehr'. Deshalb soll von alle dem, was einem insgesamt durch den Kopf gehen könnte, wenn man einen synchronisierten bzw. untertitelten Film unter die Augen und gleichzeitig an die Ohren bekommt, hier auch nur ein Bruchteil zur Sprache kommen, dabei allerdings jener, in dem sich einiges Wesentliche findet. Anders ausgedrückt soll mit diesem Beitrag nur eine Vorstellung vom Charme oder dem Zauber dieser Frage vermittelt werden. Und sollte wenigstens jetzt einem Jemand oder einer Jemandin unter den Lesern und Leserinnen aufgefallen sein, dass in diesem Beitrag die Augen nicht ohne die Ohren auftauchen, dann hänge ich daran die schon fast zum Schluss führende Bemerkung auf, dass es bei dieser regelmäßigen Paarung der wesentlichen Werkzeuge unserer sinnlichen Wahrnehmung keineswegs allein um eine rhetorische Spielerei geht, sondern um den Reflex einer der wesentlichen Besonderheiten dieser Frage, die sie zugleich so attraktiv macht. Lässt man sich nämlich auf diese Frage oder bildlich passender: mit dieser Frage ein, so tanzt man ständig auf den sprichwörtlichen zwei Hochzeiten, auf denen man laut Sprichwort nicht tanzen kann. Im gegebenen Falle muss man aber sogar auf beiden tanzen, weil es in der Natur dessen gründet, worauf die Frage, ob 'Synchronisieren oder Untertiteln' eine Antwort will. Denn das eine wie das andere hat mit Augen und Ohren gleichzeitig zu tun. Auch wenn das beim Untertiteln auf den ersten Blick gar nicht so aussehen mag. Denn auf diesen ersten Blick wendet sich ja dasjenige, was beim Untertiteln von Tonfilmen herauskommt, genauso an die Augen jeweils Zuschauender, wie es die Bilder tun. Also würde man dabei eigentlich nur auf der Hochzeit der Augen tanzen. Genauer besehen bezieht sich aber das, was da als Untertitel mit den Augen wahrgenommen wird, auf etwas, was an sich nur für die Ohren bestimmt ist, und wenn man sich nur für einen Moment vergegenwärtigt, wie jeder Mensch englische Untertitel liest, dann merkt man, dass er sie so liest, wie sie ihm in gesprochener Form an Ohr kämen. Kein Mensch, der Englisch kann, wird das Wort *laughter* so lesen, wie es geschrieben wird. Jeder wird es lesen, wie es gesprochen wird. Und gesprochen wird immer nur für die Ohren. Das Untertiteln spielt sich also auf dem Hintergrund eines Medienwechsels ab, der ganz natürlich seine Spuren hinterlässt resp. der Folgen hat, über die immer wieder gestritten wird. Zugleich verhält es sich beim Synchronisieren aber gar nicht anders. Auch hier geht es an der Oberfläche allein um ein Ohrenproblem oder wissenschaftlicher: um ein akustisches Phänomen, in der Tiefe aber sehr wohl um etwas, was die Augen teilweise geradezu schmerzlich berührt. Wenn eine beliebte Weisheit von Fernsehköchen lautet 'das Auge isst mit', dann kann man für den Film genauso gut sagen 'das Auge hört mit'.

Bei etwas genauerer Betrachtung wird man das 'Mithören' des Auges allerdings keineswegs bloß für den Film reklamieren, es gilt in gleicher Weise für die Wirklichkeit, aus der das Mithören des Auges im Übrigen auch ganz natürlich herkommt. Denn jeder weiß: Wenn jemand lächelt, hört sich alles, was er gerade sagt, anders an, als wenn er dabei die Mundwinkel nach unten zieht, ein verdrießliches oder gar grimmiges Gesicht macht oder die Augen verdreht. Nur kann man das Lächeln ja genauso wenig hören wie die herabgezogenen Mundwinkel, das grimmige Gesicht oder die verdrehten Augen. Freundlich lächelnd Gesagtes hört sich auch dann anders an, wenn experimentalphonetisch gar kein Unterschied zum Nicht-Lächelnd-Gesagten festzustellen ist. Das gilt natürlich genauso für süffisant lächelnd wie für stirnrunzelnd Gesagtes. Dieses Phänomen hat seine einfache Begründung in der Arbeit unserer Wahrnehmungsorgane. Wenn wir verschiedene Wahrnehmungssysteme unterscheiden und also von verschiedenen Sinneskanälen sprechen, über die wir etwas wahrnehmen, geschieht das im Wesentlichen aus praktischen und vor allem aus medizinischen und wissenschaftlichen Gründen. Alles Visuelle fällt in den Aufgabenbereich der Optik resp. der Augenheilkunde und alles Akustische fällt in den Aufgabenbereich der Akustik resp. Ohrenheilkunde. In der Wirklichkeit jedoch arbeiten die verschiedenen Sinne keineswegs getrennt voneinander, sondern – wenigstens beim Gesunden – immer nur zusammen und gleichzeitig. Selbst wenn wir jemanden bloß sprechen hören, machen wir uns ein Bild von ihm und sehen ihn in diesem Sinne vor uns – jeder kennt das vom Telefon –, und wenn wir jemanden nur sehen, haben wir schon eine Vorstellung davon, wie es klingen wird, wenn er den Mund aufmacht. D.h., wir hören dann schon, nicht *was* er, aber *wie* er – wenigstens ungefähr – reden wird, z.T. sogar auch schon, *was* er sagen wird. Kinder haben da ein ganz besonders feines Gespür. Am Gesicht wird klar, ob es gleich ein Donnerwetter gibt oder einen warmen Regen tröstender Worte.

In dieser Synchronie oder Gleichzeitigkeit der Arbeit der verschiedenen Wahrnehmungssysteme liegt dann aber eben auch vielleicht das Grundproblem des Synchronisierens. Weil *Gesagtes*, in diesem Falle aber auch *Gesprochenes*, und gerade zu *Sehendes* immer zusammenpassen müssen. Tun sie das nicht, spüren wir, das etwas nicht stimmt. Das gilt sogar auch für das Untertiteln, obwohl die Schrift ja völlig von den äußeren oder auch inneren Umständen, unter denen etwas gesagt wird, abstrahiert. Selbst was wir lesen, sollte zu dem passen, was wir sehen. Und dann ja auch noch hören. Denn der Originalton bleibt beim Untertiteln schließlich erhalten. Auch wenn man sich keine Illusionen machen sollte, dass Sprachunkundige wirklich etwas davon haben. Denen können höchstens die Geräusche beim Verständnis hilfreich sein. Der Rest kann eher falsche Vorstellungen wecken. Denn wenn man einen italienischen Streifen untertitelt, will das nüchterne Deutsch nur selten zu dem passen, was einem als Wortoper in die Ohren fällt.

8. Wer kann wie mit dem Unheil fertig werden?

Von hier aus kann es zu einem ziemlich kurzen Schluss kommen. Das von Kraucauer vorhergesehene Unheil, das mit dem Einzug der Sprache in die bewegten Bilder über den Film gekommen ist, will möglichst gering gehalten sein, zu ändern ist es nicht, und zwar deshalb, weil die Sprache nicht als Sprache, sondern nur in einer unglaublichen Vielfalt von Sprachen existiert. Selbst wenn 500 Millionen Menschen so gut Englisch könnten, dass sie jeden englischsprachigen Film gut verstünden, so blieben bei sechs Milliarden Erdenbürgern immer noch fünf und eine halbe Milliarden Menschen, für die ein solcher englischsprachiger Film verständlich gemacht werden müsste, wenn sie ihn so sehen sollten, dass sie hinterher sagen könnten, worum es in dem Film wirklich ging. Ob Synchronisieren oder Untertiteln ist dementsprechend keine seriöse Frage, es muss das eine oder das andere gemacht werden (Luyken 1991).

Entsprechend kann die seriöse Frage auch nur dahin gehen, wie man beides so macht, dass der angestrebte Zweck erreicht wird. Wer aber kann darauf die Antwort geben?

Für das Untertiteln wurde die Richtung angedeutet, in die eine Antwort gehen könnte: man gibt den Menschen nur zu lesen, was gesagt wird, ist also einerseits inhaltlich orientiert und andererseits pragmatisch insofern, als das zu Lesende möglichst schnell zu lesen sein sollte, weil sonst die Verweildauer des Blicks auf dem Geschriebenen zu lang ist und wesentliche Dinge im Bild nicht gesehen werden können, bevor die Szene zu Ende ist. Und Szenenende ist Untertitelende. Insofern ist Untertiteln nämlich auch eine Form des Synchronisierens.

Was die Möglichkeiten angeht, Gesagtes in einer schriftlichen Übersetzung zu komprimieren, so dass es möglichst schnell zu lesen ist, so kann tatsächlich die Linguistik im Sinne einer Linguistik der Sprache um Aufklärung ersucht werden, weil schließlich mit syntaktischen Transformationen wie Nominalisierungen oder Kompositabildungen relativ einfach Platz gespart werden kann. Für das Synchronisieren geht es aber nun überhaupt nicht um Komprimieren, sondern allein um das Imitieren, und hier kann die Linguistik der Sprachen überhaupt nichts Vernünftiges leisten, wie die Arbeit von Herbst das ganz entgegen ihrer Intention geradezu klassisch beweist (Herbst 1994). Hier kann bzw. könnte nur eine Linguistik des Sprechens weiterhelfen. Denn im Synchronisieren ist, was Herbst so überhaupt nicht erkannt hat, das Moment der Zeit enthalten, das in einer Linguistik der Sprache einfach nicht vorkommt.

Und das nun eben deswegen, weil Synchronisieren schließlich *Vergleichzeitigen* heißt, und dazu braucht es immer zweierlei Verschiedenes, das sich zur gleichen Zeit oder besser in der gleichen Zeit abspielen soll. Beginnt das Eine, muss auch das Andere beginnen, hört das Eine auf, muss auch das Andere aufhören. Die Grammatik einer Sprache ist aber in dieser Hinsicht völlig zeitlos, denn für

einen wohlgeformten Satz ist es völlig bedeutungslos, in welcher Zeit er ausgesprochen wird. Für seine Grammatikalität oder seine Akzeptabilität spielt es keine Rolle, ob für seine Aussprache genauso lange gebraucht wird, wie für einen bedeutungsgleichen Satz in einer anderen Sprache. Zeit gibt es in der Sprache allein dann, wenn gesprochen wird, und insofern gibt es Zeit auch nur im Sprechen. Das heißt sehr einfach, dass es für Fragen der Synchronisation im Bereich des Sprachlichen grundsätzlich und lediglich einer Linguistik des Sprechens bedarf, nicht der Sprache. Darüber hinaus gibt es im Phänomen des Sprechens aber, und das wiederum im Unterschied zum Phänomen der Sprache, nicht nur den Faktor resp. das Phänomen *Zeit*, sondern darüber hinaus noch viele weitere Faktoren, die uns erkennen lassen, dass gerade gesprochen wird. Denn das Sprechen ist von Natur aus und seiner Natur nach ein ganzkörperlicher Vorgang, der sich nicht allein an den Verstand, sondern immer auch an die Ohren und die Augen und zwar zur gleichen Zeit wendet. Es gilt also immer dann, wenn gesprochen wird, wenigstens Verbales, Akustisches und Visuelles zu unterscheiden resp. zu berücksichtigen. In der Schrift gibt es nichts als Verbales. Das bedeutet des Weiteren, dass bei der Analyse oder Beschreibung des Sprechens das Zusammenspiel wenigstens dieser drei verschiedenen Faktoren zu berücksichtigen ist, was etwas genauer gesagt heißt: in einer Linguistik des Sprechens spielen Stimme, Gestik, Mimik, Kinetik und artikulatorische Motorik eine Rolle, die in einer Linguistik der Sprache überhaupt nicht vorkommen. Sie sind entsprechend für jede Sprache im Einzelnen zu studieren, allerdings eben gerade nicht isoliert, sondern in ihrem Zusammenwirken. Denn sie sind für die Kommunikation ganz unzweifelhaft komplementäre Komponenten.

Damit reicht aber eben auch das Übersetzen dessen, was gesagt wird, nur in Untertiteln aus, nicht aber beim Synchronisieren. Hier ist die Übersetzung allenfalls der Rohstoff, der so zu bearbeiten ist, dass die Stimmigkeit erreicht wird, die beim Sprechen in der Originalsprache grundsätzlich vorhanden ist. Wobei bedacht werden sollte, dass in solchen Filmen normalerweise Schauspieler am Werke sind, deren Gestik, Mimik und Motorik auf entsprechende Wirkung hin ausgebildet sind. Dabei kann man als praktische Beobachtung verraten, was jeder, der sich irgendwann einmal praktisch mit dem Synchronisieren beschäftigt hat, bestätigen wird, dass nämlich Szenen, in denen der zu synchronisierende Schauspieler oder die zu synchronisierende Schauspielerin eine Bewegung mit dem Kopf, dem Arm oder dem Körper macht, immer am besten in der Zielsprache wiederzugeben sind, weil letztlich jeder ein intuitives Wissen darüber besitzt, wann beim Sprechen der Kopf zur Seite geneigt, der Arm gehoben oder der ganze Körper gestreckt wird. Natürlich immer dann, wenn in seiner Sprache gesprochen wird. Zu dem, was da allein über die Augen wahrgenommen wird, muss das passen, was gesprochen wird. Ob das zum zu Sehenden passend Gesprochene tatsächlich die linguistisch adäquate Übersetzung dessen ist, was im

Original gesagt wird, erweist sich spätestens jetzt als völlig belanglos. Für das, was von den Zuschauern als eine gute Synchronisation empfunden wird. Die wohlgemerkt die Originalsprache nicht verstehen. Übersetzer können an dieser Stelle nur den Kopf schütteln.

Aber genau deshalb kann die Antwort auf die Frage ‘Synchronisieren oder Untertiteln’ auch nicht sinnvoll von Übersetzungswissenschaftlern gegeben werden, sondern nur von Menschen, die wissen, was Menschen für ein Gesicht machen, wenn sie etwas Bestimmtes sagen wollen, wie sie dabei mit den Armen fuchteln oder mit den Beinen wippen. Und um das Fass für Linguisten der Sprache und Übersetzungswissenschaftler zum Überlaufen zu bringen: Wie an welcher Stelle einer Äußerung mit dem Kopf gewackelt, den Armen gefuchelt oder den Beinen gewippt wird, das ist von Sprachnation zu Sprachnation keineswegs notwendigerweise das Gleiche. Was schließlich bedeutet, dass es nach der Geburt des tönenden Films jeden Film nur einmal gibt. Nämlich in seiner Ausgangssprache. Alle Synchronisationen sind andere Filme.

Was Kracauer bedauerte und worin er großes Unheil für den Film sah, könnte man aber auch als Heil sehen. Insofern nämlich, als das Studium von synchronisierten Filmen zu der Erkenntnis führt, dass die Linguistik der Sprache oder der Sprachen da an ihre Grenzen stößt, wo das Sprechen als Tätigkeit anfängt. Da aber das Sprechen wissenschaftlicher Untersuchung nicht weniger wert ist als die Sprache und von dieser aus auch gar nicht vollständig zu beschreiben oder gar zu erklären ist (Coseriu 2007, 58), braucht es also zu einer Linguistik der Sprache hinzu noch etwas anderes, was solche Beschreibung und Erklärung leisten kann, und das wäre eben eine Linguistik des Sprechens, die dann wiederum auch für das Synchronisieren jene Hilfestellung leisten könnte, die die Linguistik der Sprache für das Untertiteln heute schon leisten kann.

Literatur

- Coseriu, E. 2007. *Sprachkompetenz. Grundzüge der Theorie des Sprechens*. Bearbeitet und hrsg. v. Heinrich Weber. 2. durchgesehene Auflage, Tübingen: Gunter Narr.
- Fuchs, W.J., Reitberger R. 1983. *Comics-Handbuch*. Reinbek bei Hamburg: Rowohlt.
- Herbst, T. 1994. *Linguistische Aspekte der Synchronisation von Fernsehserien. Phonetik, Textlinguistik, Übersetzungstheorie*. Tübingen: Niemeyer.
- Humboldt, W.v. 1836/1960. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Mit einem Nachw. des Verl. Faks. Dr. nach Dümmlers Orig.-Ausg. von 1836. Bonn; Hannover; Hamburg; München: Dümmler.
- Kracauer, S. 1930. “Im Westen nichts Neues”. Zum Remarque-Tonfilm. *Frankfurter Zeitung* 6.12.1939.

- Lessing, G.E. 1766/1971. *Laokoon oder über die Grenzen der Malerei und Poesie: mit beiläufigen Erläuterungen verschiedener Punkte der alten Kunstgeschichte*. Mit einem Nachwort von I. Kreuzer. Stuttgart: Reclam.
- Luyken, G.M. 1991. *Overcoming language barriers in television: dubbing and subtitling for the European audience*. With Thomas Herbst ... [et al.] Manchester: European Institute for the Media.
- Pruys, G.M. 1997. *Die Rhetorik der Filmsynchronisation. Wie ausländische Spielfilme in Deutschland zensiert, verändert und gesehen werden*. Tübingen: Gunter Narr.
- Raecke, J. 1998. Die Bedeutung der Stimme in der Fernsehwerbung – Zur Begründung einer heute möglichen Sprachwissenschaft. In: Berger, T., Raecke, J. (eds.) *Slavistische Linguistik 1997. Referate des XXIII. Konstanzer Slavistischen Arbeitstreffens. Blaubeuren. 26.-28.8.1997*. München, 144-175.
- Raecke, J. 2002. Doppelmord auf der Leinwand. Der Roman "Ubistvo s predumišljajem" als Film In: Willich-Lederbogen, H., Nohejl, R., Fischer M., Setzer, H. (eds.) *Itinera slavica. Studien zu Literatur und Kultur der Slaven. Festschrift für R.-D. Kluge z. 65. Geburtstag*. München, 215-225.
- Wikipedia: <http://de.wikipedia.org/wiki/Tonfilm>.