

Dr. Rolf Breuer

Der Mythos von der bösen Bild-Text-Schere oder: Wie es sich mit der Bild-Text-Schere gut leben lässt!

(Es handelt sich um die vorläufige Version eines Aufsatzes, die sich durchaus noch verändern kann - eine „Baustelle“ gewissermaßen. Feedback ist gerne gewünscht! Der Aufsatz darf ohne Rücksprache mit mir – außer für private Zwecke - in keiner Weise vervielfältigt werden. Copyright: siehe Fußzeile)

Die Keule

Üblicherweise benutzen Fernsehjournalisten im Berufsalltag den Begriff "Bild-Text-Schere" als Keule: Mit der gibt man einer Kollegin oder einem Kollegen voll eins auf die Mütze, weil an einer bestimmten Stelle eines Films im Text etwas gesagt wurde, was im Bild offensichtlich nicht zu sehen ist. Besonders beliebt ist der Keuleneinsatz von oben nach unten: Der Chef haut seine Redakteure oder Autoren, der ältere Kollege die jüngeren, der Fortgeschrittene die Anfänger.

Dabei fühlen sich die Keulenschwinger meist ganz toll, weil sie doch einen bösen Fehler beim anderen entdeckt haben. "Das tut man nicht!" Sie sehen den Splitter im Auge des anderen ... und übersehen den Balken im eigenen. Leider fällt das den Keulen-Opfern meist auch recht schnell auf: "Der muss grad was sagen, der macht's doch auch nicht anders!"

Was die einzelnen unter „Bild-Text-Schere“ verstehen, ist dabei nicht immer ganz dasselbe. Manche sprechen nur dann von „Bild-Text-Schere“, wenn es zwischen Bild und Text ganz offensichtlich einen Widerspruch gibt: „Ich sehe Apfel, ich höre Banane!“. Den meisten reicht es aber schon, wenn der Text etwas anderes anspricht als das Bild zeigt: „Ich sehe Apfel, ich höre etwas über den Euro.“ Andere wieder würden so etwas vielleicht auch als „Bildteppich“ bezeichnen, vorausgesetzt, das Thema „Euro“ hätte irgendwie irgendwann irgendwas mit Obst und damit mit Äpfeln zu tun. Für wieder andere Experten wären das „Sujet-Bilder“, Bilder also, die irgendwie irgendwann irgendwas mit dem Thema zu tun haben, meist aus dem Archiv stammen und sich deshalb angeblich zum „Bebildern“ des Textes eignen. Also viele Vorstellungen - und viel Durcheinander.

Unterm Strich aber ist der Begriff „Bild-Text-Schere“ im Alltag der Fernsehjournalisten ein Synonym für handwerkliches Versagen. Es dominiert die Vorstellung: Die Bild-Text-Schere ist etwas durch und durch Negatives, „Böses“! Im Folgenden werde ich aber ausführen, dass das nicht einmal die halbe Wahrheit ist - es ist noch viel weniger. Es ist nichts weiter als ein Mythos, ein Ammenmärchen. Denn bei genauerem Hinsehen kommt man zu der Einsicht:

Es gibt keinen Informationsfilm ohne Bild-Text-Schere.

Mehr noch:

- **Sie ist prinzipiell unvermeidlich,**
- **ästhetisch regelrecht notwendig**
- **und hat sogar kreatives Potential!**

Aber sie kann leider auch heftig destruktiv sein.

Und das ist das große Problem.

Der Anfang samt hübschen Effekten

Wie es aussieht, hat alles in den 70er Jahren angefangen mit Bernward Wember und seinen Untersuchungen (in Film und Buch): "Wie informiert das Fernsehen?" In diesem Zusammenhang prägte er wohl auch den Begriff der "Bild-Text-Schere" - wenn also die Bild-Information und die Text-Information an einer Stelle eines Fernsehfilmes heftig auseinander liegen oder auseinander driften.

Folge davon ist häufig, dass das Publikum in der betreffenden Passage weniger versteht und dass oft nur Bruchstücke des Ganzen im Gehirn ankommen: ein bisschen Text und etwas mehr Bild - oder gelegentlich auch umgekehrt. Eine Beobachtung aus meinen Seminaren: Manchmal kommt es sogar vor, dass man konzentriert auf den Bildschirm geschaut hat, aber das Bild für kurze Zeit nicht gesehen hat! Ein hübscher Effekt: Man guckt Fernsehen und sieht nichts. (In diesen Fällen war dann entweder der Text ausnahmsweise mal stärker als das Bild - zum Beispiel stark emotional geladen - oder die vorangegangene Passage machte heftige Denkarbeit nötig. Letzteres führte dann auch dazu, dass für einen Moment weder Text noch Bild erinnert wurde.)

Eine andere Folge kann sein, dass - beabsichtigt oder unbeabsichtigt - Bilder einen anderen Sinn bekommen als sie eigentlich haben. Wenn zum Beispiel in einer Gerichtsreportage ein unbekannter Zeuge Z zu sehen ist, die Rede aber von den Verbrechen des nicht zu sehenden (und ebenso unbekannt) Angeklagten A ist: Da kriegt der Zeuge plötzlich ein gefährliches Image!

Bernward Wember hat damals auch gleich ausprobiert, was man an solchen Stellen denn tun könnte: Wo abstrakter Text mit Hintergrund-Informationen notwendig war, ersetzte er z.B. die Großeinstellung "marschierende Soldatenstiefel" durch eine ruhige Landschaftstotale. Effekt: In einer Zuschauer-Vergleichsgruppe hatten viel mehr Personen verstanden, was gemeint war. Die Schere war nicht mehr so groß, weil die ruhige Landschaft weniger aufreizend, dramatisch und emotional war als die Einstellung „Soldatenstiefel“.

Sein zweiter (Haupt-)Vorschlag war, an solchen Stellen mit Grafiken zu arbeiten - aber das überstrapazierte er als Lösung dann doch etwas und machte sich damit leicht angreifbar. Schade eigentlich, denn im Prinzip hatte er Recht: Starke Textreize brauchen schwache Bildreize, starke Bildreize verlangen schwache Textreize.

Das ist natürlich im Moment noch sehr abstrakt und grob vereinfacht: Welche Bild- oder Textreize sind denn "stark", welche "schwach"? Gilt das pro Bild und Textteil absolut - oder relativ, in Abhängigkeit voneinander? Oder beides?

Das destruktive Potential der Bild-Text-Schere

Jedenfalls kann man festhalten: Bild-Text-Scheren können bei Informationsfilmen ein Problem sein - wenn ein starker Bildreiz nämlich dazu führt, dass in der betreffenden Passage die gesprochene Zusatzinformation bei vielen Zuschauern gar nicht oder nur in Bruchstücken ankommt. Da hat die Bild-Text-Schere dann regelrecht destruktives Potential.

Das deckt sich mit meinen Erfahrungen aus über 20 Jahren Trainertätigkeit mit weit über 2.000 Analysen von Magazinbeiträgen. Dabei fanden sich in fast jedem Film bei einer genauen Feinanalyse Textstellen, die einfach nicht gehört worden waren - von allen oder von einem großen Teil der anwesenden Teilnehmer.

Glück hatte der Autor, wenn es nicht gerade Schlüsselwörter oder Schlüsselpassagen waren. Und er hatte Pech, wenn es sich doch um wichtige Informationen handelte. Dann war der Sinn einer Passage gar nicht zu verstehen oder nur verfälscht bis hin zum Gegenteil dessen, was gemeint war. Und wenn der Satz ein Schlüsselsatz für eine bestimmte Passage war, war die Passage "tot"; und wenn die Passage eine Schlüsselpassage für den gesamten Film war, dann ... Öfters hatte sogar ein Teil der Gruppe den ganzen Film völlig anders verstanden als er gemeint war. Fast immer ließ sich das auf solche "verlorenen" Textpassagen zurückführen.

Erkannte und unerkannte Bild-Text-Scheren

Es ist dabei aber regelrecht paradox: Die **erkannten** Bild-Text-Scheren (s.o. "Die Keule") sind gar nicht so sehr das Problem. Denn wer eine Bild-Text-Schere beim ersten Sehen erkannt hat, der hat an dieser Stelle den Text wenigstens ungefähr verstanden. Er hat zum Beispiel erkannt, dass Herr Z zu sehen ist, aber der Name von Herrn A zu hören ist. Und der ist blöderweise der Verbrecher, Herr Z aber nur der Zeuge.

In einem solchen Fall ist die offensichtliche und gleich erkannte Bild-Text-Schere dann sogar ganz nützlich: Bei der Abnahme kann ein solcher Fehler vielleicht noch korrigiert werden. Und manchmal sind solche offensichtlichen Bild-Text-Scheren ja auch ganz unterhaltsam, weil freiwillig-unfreiwillig komisch: Wenn etwa ein Schwein zu sehen ist, der Text aber vom Landwirtschaftsminister spricht.

Kleiner Exkurs: Solche erkennbaren Bild-Text-Scheren kann man bewusst ausnutzen - als Anstoß zum Nachdenken, als Provokation oder etwa für eine Glosse oder Satire. Man baut darauf, dass mit einem bestimmten Bild (Person oder Sache) eine bestimmte Vorstellung bei den meisten Zuschauern verbunden ist. Der Text gibt dem Bild aber einen neuen Sinn (in der Art "Schwein / Landwirtschaftsminister") und führt damit - hoffentlich - zu einer interessanten neuen Aussage (Satiren á la Monitor gehören etwa dazu). So etwas wäre dann eine **gewollte** oder **kontrollierte** Bild-Text-Schere, die geradezu darauf baut, erkannt zu werden. Sie kann durchaus auch in ernsthaften Zusammenhängen benutzt werden.

Aber zurück zum Erkennen oder Nichterkennen: Viel gefährlicher als die erkannten sind die **unerkannten** Bild-Text-Scheren - die man nämlich beim ersten Sehen gar nicht entdeckt, sondern höchstens bei einer nachträglichen Feinanalyse. Sie sorgen unangenehmerweise dafür, dass Text-Informationen einfach im Nirwana verschwinden - ohne dass es den Betreffenden bewusst ist. Das ist ein Kerndilemma: Wenn ein starker Bildreiz (Umschnitt oder Unerwartetes) den Text zertrümmert, kriegt das der Zuschauer in der Regel nicht mit: Seinem Gehirn reicht an der Stelle das starke Bild; Text hat da für einen Moment einfach nicht stattgefunden. Und deswegen kann (etwa bei einer Abnahme) diese Stelle häufig auch nicht als Problemstelle erkannt werden. Aber wie kommt es dazu?

Gehirnarbeit nach Regeln

Wenn unser Gehirn einem Film bzw. Fernsehbeitrag zu folgen versucht, gibt es dabei - vereinfacht gesehen - drei wichtige Regeln:

Erstens kann unser Gehirn in einem Moment immer nur einen Gedanken denken: um etwa ein Bild oder einen Text zu verstehen bzw. zu entschlüsseln. Mehrere Gedanken können nur hintereinander bzw. hin- und herspringend gedacht werden. Daraus folgt beispielsweise: Man kann nicht zwei Geschichten im gleichen Moment verfolgen: eine interessante Bildgeschichte und zugleich eine völlig andere Textgeschichte!

Zweitens hat das Gehirn immer nur eine bestimmte Menge Aufmerksamkeit zur Verfügung.

Da reicht schon alleine ein Umschnitt von einer Einstellung auf eine völlig andere, um nach dem Umschnitt nur durch das optische Ereignis für ein bis zwei/drei Sekunden voll ausgelastet zu sein! Text, der an dieser Stelle nichts mit dem neuen Bild zu tun hat, geht sehr häufig in diesen ein bis zwei Sekunden verloren. Er wird zwar physikalisch gehört, aber nicht mit Sinn wahrgenommen. Viele Seminarteilnehmer können sich nach dem ersten Sehen eines Films ("Zuschauersehen") nicht erinnern, an solchen Stellen bestimmte Wörter gehört zu haben. Der Bildreiz ist zu hoch, um noch einen starken Textreiz verdauen zu können. Nach diesen ein bis zwei/drei Sekunden lässt dann übrigens der Bildreiz für das Zuschauer-Gehirn mehr und mehr nach, so dass wieder Kapazität zum sinnvollen Zuhören frei ist - ein wichtiger Effekt für die Praxis.

Die gleiche Überlastung geschieht übrigens nicht nur bei den "gefährlichen" Umschnitten, sondern auch dann, wenn sich innerhalb einer Einstellung der Bildinhalt überraschend und bilddominierend ändert. Das kann mal nur der breite Rücken eines Kameramannes sein, der ein wichtiges Geschehen verdeckt, oder ein unerwartetes dramatisches Ereignis im Bild. In allen Fällen braucht das Gehirn eine gewisse Zeit, um den neuen unerwarteten Bildreiz zu entschlüsseln und ggfs. zu verdauen (bei hohem emotionalen Inhalt etwa).

Anders gesagt: Unser Gehirn kann nicht in einem Moment einen starken Bildreiz sinnvoll verarbeiten und zusätzlich noch einen starken Textreiz, der nichts mit dem Bild zu tun hat. Die meisten Gehirne entscheiden sich in einem solchen Fall übrigens für das Bild (etwa 80% der Zuschauer), die übrigen können vielleicht gerade noch zuhören, ohne zu stark vom Bild abgelenkt zu werden.

Drittens ist unser Gehirn immer bestrebt, alle vorhandenen Sinneseindrücke zu einem "Ganzen" zusammenzufügen. Deswegen versucht es beim Film-Sehen immer, Gesehenes mit Gehörtem zu verbinden. Wenn das klappt, "läuft" ein Film an der Stelle gut. Entstehen aber Irritationen (weil der Sinn eines Bildes nicht klar ist und / oder weil der Text nur in Bruchstücken oder gar nicht angekommen ist), dann beginnt das Gehirn damit, durch Nachdenken Lösungen zu finden - eben den Sinn des Ganzen. In der Phase des Nachdenkens aber kann es dann wieder nicht richtig zusehen und / oder zuhören - und so steigt das Risiko, dass an diesen Nachdenk-Stellen Informationen verloren gehen. Übrigens können da auch Text-Pausen Abhilfe schaffen - als "Verdauungspausen" quasi.

Soweit also einige Gedanken zum destruktiven Potential der Bild-Text-Schere und ihren Hintergründen in unserem Gehirn. Aber wie lassen sich diese negativen Effekte vermeiden? Der Begriff der "Schere" sagt noch nicht viel dazu - außer vielleicht: "Scheren vermeiden". Also weg damit?

In den vielen Jahren als Trainer hat sich allerdings eine völlig entgegengesetzte Einsicht verfestigt, die nur auf den ersten Blick überraschend ist:

Die Bild-Text-Schere ist unvermeidlich bzw. strukturell notwendig

Informationsfilme (Nachrichten, aktuelle und andere Magazine, Dokumentationen ...) zeichnen sich dadurch aus, dass **immer** wesentliche Teile der Gesamt-Information über gesprochenen Text vermittelt werden müssen. Zwangsläufig bedeutet das: Es muss (viele) Stellen geben, an denen Bild und Text auseinander liegen. Es sei denn, man würde nur O-Töne im ON nehmen oder den Text eins zu eins als Schrift durchlaufen lassen ohne jegliches Bildgeschehen.

Das heißt also: Bild-Text-Scheren in Informationsfilmen sind prinzipiell unvermeidlich und strukturell regelrecht notwendig. Da kommt man nicht drumrum. Man kann nur mit der Schere leben - und sie möglichst geschickt einsetzen. So gesehen ist die Bild-Text-Schere also ein ganz zentrales Werkzeug der Fernsehjournalisten.

Aber nicht nur das:

Die Bild-Text-Schere ist ästhetisch notwendig

Die Bild-Text-Schere ist nicht nur strukturell notwendig, nein, unser Gehirn lechzt gewissermaßen danach! Nur lesen oder sehen, was ich höre, wäre auf Dauer ja stinklangweilig. Ich möchte gerne was Interessantes sehen und hören - also starke oder wichtige Reize auf beiden Ebenen. Aber (s.o.) bitte nicht gleichzeitig - das schafft erfahrungsgemäß kein Gehirn. Sondern nacheinander: Zum Beispiel erst ein starker Bildreiz und danach ein starker Textreiz.

Zur Illustration nehmen wir mal an, man arbeitet an einem Film über Hüte.

Einstellung: "Hut groß in der Vitrine, langsame Rückfahrt, 5 sec." (zugegeben, das ist nicht besonders fantasievoll - aber es soll ja hier nur Prinzipielles illustrieren). Der Text beginnt jedes Mal nach ca. 0,5 sec.

Text 1: "Das ist ein Hut." Gaaanz toll! "Sehe ich doch!" - sagt sich vermutlich jedes Zuschauer-Gehirn und findet, dass man es für blöd hält.

Text 2: "Der Hut hat Konrad Adenauer gehört!" Ist schon besser, ich bekomme eine Information, die über das hinausgeht, was ich sehe. Das findet mein Gehirn schon interessanter und ist ganz zufrieden.

Text 3: "Der hat Adenauer mal das Leben gerettet!" Das ist ja dann schon fast spannend!

In den Versionen 2 und 3 stellt der Text als erstes eine Verbindung zum (neuen) Bild her und gibt danach eine Information, die über das Bild hinausgeht. Gleichzeitig nimmt der Bildreiz allmählich ab.

Aber trotzdem: Im hinteren Teil der Einstellung haben wir nach wie vor eine Bild-Text-Schere! Ich sehe nämlich "Hut", aber ich höre "hat Konrad Adenauer gehört" bzw. "hat Adenauer das Leben gerettet". Nur: In dieser Weise funktioniert es offenbar, mein Gehirn findet diese Bild-Text-Schere sinnvoll und interessant - in anderen Worten: schön. Ich bezeichne das deshalb als ein "ästhetisches" Bedürfnis nach einer Bild-Text-Schere.

"Schön" ist eine solche Passage mit Bild und Text nur, wenn ich erstens das Bild und zweitens den Text verstehe und drittens beides nicht identisch ist. (Außer in Notfällen, wenn etwa ein komplizierter anstrengender Gesetzestext in Wort und Schrift zitiert wird).

Alltagsmodell "Dia-Vortrag"

In der Praxis hat man es aber nicht nur mit einer einzigen Einstellung ("Hut") zu tun, sondern mit einer Folge von Einstellungen. Und das bedeutet: Immer wieder kommt ein Umschnitt (ein neues Bild) und treibt den Bildreiz binnen Sekundenbruchteilen stark in die Höhe. Oft verliert man als Zuschauer die Textinformation, die sofort nach dem Umschnitt gesprochen wurde (in den ersten ein bis zwei/drei Sekunden meist). Wie soll man damit umgehen?

Als Modell dient eine kleine Geschichte aus dem Alltag: Herr und Frau Müller zeigen ihren Freunden ihre Urlaubs-Dias. Wenn ein neues Dia erscheint (Umschnitt), reden die Müllers sofort los, etwa in der Art: "Dahinten ist unser Haus!" oder "Guckt mal, das war kurz vor dem Unwetter!" Das heißt, ganz intuitiv nehmen die Müllers den hohen Bildreiz des neuen Bildes sehr ernst und platzieren einen Text dazu, der das neue Bild verstehen hilft. Dabei sprechen sie entweder die Hauptreize des Bildes an oder führen die Aufmerksamkeit ihrer Gäste zu einem Detail, das aus ihrer Müller-Sicht jetzt das Wichtigste ist. Wenn das Dia dann länger steht, erzählen sie noch Zusatzgeschichten dazu, also Text, der über das Bild hinaus geht. Das funktioniert, weil der Bildreiz ja immer schwächer wird: erstens haben die Müllers das Bild wenigstens zum Teil erklärt und zweitens ändert sich ja auch nichts mehr

am Bild. Wenn's gut geht, finden sie im Lauf der Zusatzgeschichten sogar noch mal eine Verbindung zum ganzen Dia oder zu einem Detail, das auf den ersten Blick keine Rolle spielte.

Es gibt aber auch noch eine zweite Variante. Auf der Leinwand erscheint ein romantischer Sonnenuntergang. Noch beim Dia davor hatte Frau Müller gesagt: "Jetzt kommt mein Lieblings-Dia!" Weil das Bild sooo schön ist und die Müllers ganz ergriffen sind, sagen sie (nach dem "Umschnitt") ein paar Sekunden einfach gar nichts, lassen den Sonnenuntergang einfach nur wirken. Und dann kommt Herr Müller mit der Bemerkung: "Ist doch wirklich schön, oder?", die Gäste bestätigen es, und Herr Müller erzählt stolz, wo sie bei dem Foto gestanden haben und wie schwierig es mit dem Gegenlicht war und so weiter.

So geht man also im Alltag mit Umschnitten um: Erst eng am Bild texten (auftauchende Fragen beantworten) oder das Bild kurz ohne Text wirken lassen, damit jeweils der starke Anfangsreiz abnimmt. Danach kommen dann die Informationen, die über das Bild hinaus gehen.

Weit verbreitet: Das Irgendwannmal-Prinzip

Im Fernsehen ist aber alles anders. Da vergessen die vielen Müllers als Redakteure oder Autoren, wie sie im Alltag "aus dem Bauch heraus" mit neuen Bildern und starken Bildreizen umgehen. Stattdessen hat sich unter den Profis eine andere Regel breit gemacht: das "Irgendwannmal-Prinzip".

Übliches Muster: In 30 Sekunden berührt der Text irgendwann mal das Bild - einmal, zweimal oder (bei Super-Profis) vielleicht sogar dreimal! Da bildet der Text also "irgendwannmal" mit dem Bild eine Art Einheit und geht zufällig auf die eine oder andere Frage ein, die das Bild stellt. Die entscheidende Rolle spielt der Text: Wenn er es erlaubt, wird eine Verknüpfung zum Bild hergestellt - sonst nicht.

Nun ist das sicherlich besser als gar nichts, das muss man diesen Profis schon lassen. Aber eine richtig nützliche Strategie ist das (s.o.) leider auch nicht. Zu oft zertrümmern bei diesem Irgendwannmal-Prinzip die starken Bildreize (Umschnitte, Bildinhalte) nämlich dem Zuschauer die Textinformation, nur Bruchstücke kommen an - welche, ist relativ unberechenbar.

Und vor allem werde ich als Zuschauer bei diesem Vorgehen mit zwei sehr widersprüchlichen Strategien konfrontiert: Umschnitt, ein neues Bild, der Text ist eng am Bild. Damit sagt mir der Macher: "Was Du jetzt hörst, ist auch das, was Du siehst!" Wenige Sekunden später, wieder ein neues Bild, der Text hat nichts damit zu tun. Jetzt ist es die entgegengesetzte Botschaft: "Was Du jetzt hörst, hat überhaupt nichts mit dem zu tun, was Du gerade siehst!" Da blicke einer nur durch!

Absolute, inhaltliche und relative Reize beim Fernsehen

Es sind eben eine ganze Menge an "Reiz-Faktoren", die in einer bestimmten Passage eines Informationsfilmes zusammen wirken, auf den Zuschauer einstürzen und vom Autor möglichst beherrscht werden müssen:

absolute Bildreize: Reize, die von der äußerlichen Bildfolge einer Sequenz ausgehen (kurze Schnitte - lange Schnitte, Naheinstellungen - Totalen, Schwenks und Zooms - Standbilder, Unschärfen - Schärfen, Farbkontraste, Hell-Dunkel-Kontrast). Anders ausgedrückt: Es geht dabei darum, wie schwierig es für den Zuhörer ist, die Bildsequenz beim ersten Sehen zu entschlüsseln - unabhängig vom Text und so, wie es vom Autor gemeint war .

inhaltliche Bildreize: Reize, die vom verstandenen Inhalt einer Bildpassage ausgehen (Überraschungen, Irritationen, Emotionen, neues Wissen, Nachdenken) - unabhängig vom Text

relative Bildreize: Reize, die vom Bild ausgehen in Verbindung mit dem, was vom Text verstanden wurde

absolute Textreize: Reize, die vom Satzbau und der Wortwahl einer Textpassage ausgehen sowie von der Präsentation durch den Sprecher (abstrakt - konkret, lang - kurz, verschachtelt - linear, kompliziert - einfach, laut - leise, unklar - klar, falsch betont - richtig betont). Anders ausgedrückt: Es geht dabei darum, wie schwierig es für den Zuhörer ist, die Textpassage beim ersten Hören zu verstehen - so, wie es vom Autor gemeint war und zunächst noch unabhängig vom Bild.

inhaltliche Textreize: Reize, die vom verstandenen Inhalt einer Textpassage ausgehen (Überraschungen, Irritationen, Emotionen, neues Wissen, Nachdenken) - unabhängig vom Bild

relative Textreize: Reize, die vom Text ausgehen in Verbindung mit dem, was vom Bild verstanden wurde

Nicht zu vergessen: O-Töne, Atmo und Musik

Und dann sind da ja auch noch drei weitere "Qualitäten" auf der akustischen Ebene: O-Töne, Atmo und Musik. Sie selbst senden auch jeweils absolute, inhaltliche und relative Reize aus und können sich störend oder verstärkend auf den "Sinn des Ganzen" auswirken!

Typische Probleme sind dabei zum Beispiel:

- Musik mit starken Akzenten oder gar Gesang, was es an einzelnen Stellen immer wieder schwer macht, den Text zu verstehen. (Wer macht eigentlich zu Hause extra die Stereoanlage an, wenn er seinem Partner etwas Wichtiges erzählen will?)
- Musik, die stilistisch-emotional nicht zur gemeinten Aussage einer Passage passt und damit ungewollt irritiert
- Atmo, die zu laut ist und damit die Verständlichkeit des Textes erschwert
- Atmo, die etwas anderes hörbar macht, als im Moment zu sehen ist (auch: sprechende Menschen) und damit ungewollt irritiert
- O-Töne, die in Kommentar-Text hineinlaufen und die Verständlichkeit damit behindern
- O-Töne, die wegen schwerer Verständlichkeit oder wegen dramatischer Inhalte "nachgearbeitet" werden müssen und damit die Aufmerksamkeit für den Kommentar-Text behindern

Wenn man also bedenkt, wie viele Faktoren einen Einfluss auf Gelingen oder Nicht-Gelingen beim Texten zum Bild haben, bleibt nur ein Schluss: Es ist einfach ein verdammt kompliziertes Medium, das Fernsehen!

Besser dem Reizstärke-Prinzip folgen!

Statt diese vielen Reiz-Faktoren mehr oder weniger zufällig ihre ganze Störkraft entfalten zu lassen ("Irgendwannmal-Prinzip"), halte ich es für wesentlich sinnvoller, nach dem "**Reizstärke-Prinzip**" zu verfahren - so wie im Alltags-Dia-Vortrag oder wie im Beispiel "Hut".

Das bedeutet: Als Texter muss ich herausfühlen, wann die Bildreize so stark sind, dass ich mit dem Text ganz nah an sie herangehen muss. Man kann dann im Text zum Beispiel offene Fragen des Bildes ansprechen, erläutern und so dem Bild Reizstärke nehmen. Das kann aber auch mal durch eine Textpause geschehen (wie beim Sonnenuntergang im "Dia-Vortrag"). Auch so verliert das Bild an Reizstärke, und danach kann man mit dem Text über das Bild hinaus gehen.

Das Reizstärke-Prinzip bedeutet also:

Das Bild führt den Text, der Text die Information

Blöderweise gibt es kein elektronisches Messgerät, mit dem man bestimmen könnte, welcher Bildreiz denn welchen Textreiz verträgt. So ist denn ein jeder auf sich selbst angewiesen - und auf andere Personen, mit denen er sein eigenes Urteil vergleichen kann.

Hauptstrategie: erstens sich beim Texten von den starken Bildreizen (Umschnitte, Veränderungen im Bild) führen und sogar inspirieren lassen (s.u. "kreatives Potential") und zweitens ganz nah an der Alltagssprache formulieren. Hauptkontroll-Methode (s.o. "Dia-Vortrag"): den vorgesehenen Text laut zu den Bildern sprechen und dabei aufs Bild sehen. Da spürt man dann mehr oder weniger deutlich, ob es so geht, ob es nicht so richtig geht - oder ob es gar nicht geht.

Wer das Bild nicht ernst nimmt, dem erschlägt es den Text.

Also konkret: Zum Beispiel nach Umschnitten etwa eine halbe Sekunde das Bild erst einmal frei stehen lassen. Dann mit dem Text beginnen, der nah ran geht ans Bild und auf es eingeht. Dabei verringert sich der Bildreiz, und man kann dann zunehmend Text platzieren, der über das Bild hinausgeht - aber er muss trotzdem erzählerisch einfach sein.

Damit aber nicht nach zwei, drei Sekunden ein neuer hoher Bildreiz (z. B. Umschnitt) schon wieder den Text "zwingt", empfiehlt es sich, Einstellungen ruhig mal etwas länger stehen zu lassen: also lieber dreimal sechs Sekunden als sechsmal drei Sekunden. Deutsche Sätze sind eben üblicherweise nicht nur zweieinhalb Sekunden lang.

Natürlich ist damit nicht gemeint, streng immer nur Sechs-Sekunden-Einstellungen zu benutzen - Abwechslung muss schon sein. Aber man wundert sich oft, wie lang eine Einstellung sein kann, ohne langweilig zu werden, wenn sie sinnvoll und verständlich mit einem interessanten Text verbunden ist. Meinem (Zuschauer-)Gehirn ist es egal, ob eine Stelle interessant ist nur wegen des Bildes, nur wegen des Textes oder wegen Text und Bild!

Der nächste Einwand könnte sein, dass Sequenzen fürchterlich hölzern wirken können, wenn jeder Umschnitt frei von Text ist. Stimmt. Aber lieber hölzern als unverständlich. Und außerdem gibt es ja noch die eleganten Lösungen, bei denen der Text das Bild trotzdem ernst nimmt. Zum Beispiel legt man die Sprechpause zwischen Haupt- und Nebensatz oder zwischen zwei Satzteilen auf den Umschnitt und textet dann so, dass nach dem Umschnitt als erstes ein enger Kontakt zum neuen Bild entsteht. Aber dann muss das Timing auch sehr präzise sein, und kleine Abweichungen von nur 0,5 Sekunden können in manchen Fällen böse Folgen haben (das Schlüsselwort kann zum Beispiel verloren gehen).

Text-Pausen helfen Verdauen!

Und noch eine wichtige Hilfe, vor allem wenn der Text etwas anspruchsvoller ist (vom Inhalt her - nicht von der Form bitte!): genügend Text-Pausen einplanen! Sie helfen den Zuschauern als "Verdauungspausen" - aber nur, wenn sie sinnvoll sind. Und das sind sie, wenn es in dem Moment zum Beispiel etwas Wichtiges zu denken oder sehen oder zu hören gibt (vielleicht sogar alles zusammen). "Wichtig" bedeutet: wichtig für den Sinn, für die Aussage dieser Passage. "Unwichtige" Text-Pausen (deren Sinn ich nicht verstehe) können dagegen irritieren, werfen neue Fragen auf, die nicht beantwortet werden und deshalb auch schnell vom Wesentlichen ablenken können.

Texten zum Bild: Erzählen statt Quälen!

Vor diesem Hintergrund ist eines bisher sicher deutlich geworden: Die wichtigste handwerkliche Methode, mit der man die Fehlerquellen beim "Texten zum Bild" verringern kann, ist die "innere Erzählhaltung". Also zum Bild in etwa so texten, wie man auch beim Dia-Vortrag zu Hause sprechen würde: eng an den starken Bildreizen entlang und sprachlich wie beim Abendbrot oder im Café. Diese "innere Erzählhaltung" hilft beim Entwerfen des Textes, beim ersten Gegenlesen zum laufenden Bild, beim Korrigieren und beim Finden von Alternativen (s.u. "kreatives Potential der Bild-Text-Schere").

Dabei braucht niemand etwas völlig Neues zu lernen: Denn über die "innere Erzählhaltung" des Alltags verfügt jeder, hat sie seit vielen Jahren tagtäglich und intensiv trainiert, weiß, wie man Informationen erfolgreich an den Mann oder die Frau bringt. Sicher, ein paar wenige Prozent der eigenen Alltagssprache wird man dann für's Fernsehen noch etwas nachbessern müssen (Slang, Kraftausdrücke usw.), aber im Übrigen funktioniert das sehr gut, weil auch die Zuschauer darin trainiert sind, solche einfache Sprache beim ersten Hören zu verstehen.

Aber seltsamerweise meinen viele Menschen, dass "Erzählen wie im Alltag" bei Fernsehbeiträgen nichts zu suchen hat. Und deshalb haben sie sich eine gestelzte, komplizierte Kunst-Haltung angewöhnt, demonstrieren, was für tolle „**Schlau-Spieler**“ sie doch sind. Sie greifen dabei auf (scheinbar) erfolgreiche Modelle nicht nur aus dem Fernsehen zurück und kopieren doch nur die Fehler dieser Modelle.

Die Bild-Text-Schere hat kreatives Potential

Wenn man sich auf die Erzähl-Methode beim "Texten zum Bild" einige Zeit eingelassen hat, wird man recht bald mit einer ganz unerwarteten und positiven Eigenschaft der Bild-Text-Schere konfrontiert: mit ihrem kreativen Potential.

Oft hat man eine bestimmte Text-Idee im Kopf, aber leider verträgt sie sich nicht mit den vorhandenen Bildreizen an dieser Stelle. Die Information muss aber hier gegeben werden, und ein neuer Schnitt ist aus Zeitgründen nicht möglich. Meist ist man dann ja gewissermaßen vernagelt im Gehirn, klammert sich an der ersten Text-Idee fest - und kommt nicht weiter. Kein Kollege weit und breit, der (wegen Nicht-Beteiligung) nicht so vernagelt ist.

Was man dann probieren kann, ist, das Bild mehrmals durchlaufen zu lassen und einfach "wie zu Hause beim Dia-Vortrag oder beim Abendessen" ein paar Varianten Text gegenzusprechen - ohne groß zu überlegen, ohne Text vorzuformulieren, eher aus dem Bauch heraus. Und zwar so, dass das Bild oder Teile von ihm (direkt oder indirekt) gleich am Anfang im Sprech-Text vorkommen und dass danach in ganz einfacher alltagsnaher Weise weitergesprochen wird. Wo es nötig ist oder wo es

sich intuitiv anbietet, werden dann auch zwischen den Umschnitten hohe Bildreize im Gesprochenen aufgenommen. Man sollte dabei sehr, sehr alltagsnah sprechen - bloß kein Schriftdeutsch, denn sonst kann man sich kaum vom schon vorhandenen Text lösen.

Spätestens nach vier, fünf Versuchen (Obergrenze: zwei Minuten!) hat man dann oft eine neue, einfachere und meist sogar kürzere Lösung - eine Lösung, zu der einen das Bild und (wenn man so will) die unvermeidliche Bild-Text-Schere angeregt haben. Bei den Übungen in meinen Seminaren ist es immer wieder verblüffend, wie kreativ die so gefundenen Lösungen oft sind. Allerdings mit einer Einschränkung: Wenn man alleine ist, ist die kreative Ausbeute nicht so hoch wie dann, wenn weitere Personen beteiligt sind und sich gegenseitig zu neuen Ideen anstoßen.

Das Fazit

Fassen wir also noch einmal zusammen:

Er ist weit verbreitet, der Mythos von der „bösen“ Bild-Text-Schere.

Aber er beschreibt nur einen winzigen Teil des Problems „Texten zum Bild“. Und er gibt kaum praktische Hilfestellungen zur Lösung. Der Wirklichkeit näher kommt man, wenn man von einer anderen Sicht der Dinge ausgeht:

Es gibt keinen Informationsfilm ohne Bild-Text-Schere.

Mehr noch: Sie ist prinzipiell unvermeidlich, strukturell und ästhetisch regelrecht notwendig und hat sogar kreatives Potential!

Aber sie kann leider auch heftig destruktiv sein.

Um das zu verhindern, sollte man aus der "inneren Erzählhaltung" heraus am Bild entlang texten: Das Bild führt den Text, der Text die Information.

Aber nach aller Erfahrung hilft auch das nicht, den Perfektions-Olymp der Informationsfilme zu erklimmen: 100%-Lösungen gibt es nicht, Fehler kommen auch bei den berühmtesten Autoren immer wieder vor - zu schwierig ist eine klare, verständliche, interessante Gestaltung von Bild und Text für große Zuschauergruppen. Und trotzdem lohnen sich die beschriebenen Anstrengungen, denn mit jeder einfacheren Passage (erzählerischer Text eng am Bild) gewinne ich wieder eine Menge Zuschauer dazu, die das verstehen, was ich ihnen nahe bringen will.

Es ist einfach ein verdammt kompliziertes Medium, das Fernsehen.

Aber nur Mut:

Schlechte Beiträge gibt es genug - es kann nur besser werden!