

LEITFADEN III: Kognitive Grundlagen multimodaler
Praxis

MANUAL III: Cognitive foundations of multimodal
practice

PD Dr. phil. habil. Ilona Schulze

Zweisprachiger Leitfaden zu den kognitiven Grundlagen multimodaler Praxis im öffentlichen Raum
Bilingual manual on the cognitive foundations multimodal practice in public space

Juli – July 2020

INHALTSVERZEICHNIS/CONTENTS

1	EINLEITUNG	3
2	TEXT UND BILD	4
2.1	ANEIGNUNG VON KOMMUNIKATEN:	5
2.2	UNTERSCHIEDUNG VON WAHRNEHMUNGSARTEN	11
2.3	TEXT-BILD-KOMBINATIONEN.....	12
3	SEHSCHULUNG	15
4	BILDER, FARBEN, SPRACHE - DIE AUSSAGEKRAFT MULTIMODALER DATEN	20
4.1	KOMMUNIKATE ALS SPIEGEL SOZIALER STRUKTUREN UND TRADITIONEN	20
4.2	SPRACHFREIE PRÄSENTATION LOKALER LEGENDEN UND TRADITIONEN	21
4.3	AGGREGATE AUS TEXT UND ARCHITEKTUR.....	23
4.4	FUNKTION UND BEDEUTUNG NON- UND PARAVERBALER ELEMENTE.....	25
4.5	ZUSAMMENSPIEL MEHRERER KOMponentEN IN KOMPLEXEN KOMMUNIKATEN	27
4.6	NONVERBALE BETONUNG EINES THEMAS	30
5	BILDER, FARBEN, SPRACHE – DIE GRENZEN SEMIOTISCHER SYSTEME	32
5.1	GRENZEN DER MODALITÄT FARBE	32
5.2	BEDEUTUNGSVARIANZ VON BILDERN.....	34
5.3	DIE GRENZEN VERSCHRIFTETER SPRACHE.....	39
6	INTRODUCTION	41
7	TEXT AND IMAGE	42
7.1	ACQUISITION OF MULTIMODAL AGGREGATS:.....	43
7.2	DISTINGUISHING WAYS OF PERCEPTION.....	49
7.3	TEXT-IMAGE COMBINATIONS	50
8	VISUAL TRAINING	53
9	IMAGE, COLOUR, LANGUAGE – THE INFORMATIONAL LOAD OF MULTIMODAL DATA	59
9.1	MULTIMODAL AGGREGATES AS MIRROR OF SOCIAL STRUCTURES AND TRADITIONS.....	59
9.2	LANGUAGE FREE PRESENTATION OF LOCAL MYTHS AND TRADITIONS	60
9.3	TEXT-ARCHITECTURE AGGREGATES.....	62
9.4	FUNCTION AND MEANING OF NON AND PARAVERBAL ELEMENTS	64
9.5	THE INTERACTION OF COMPONENTS IN COMPLEX MULTIMODAL AGGREGATES	66
9.6	NONVERBAL EMPHASIS OF A TOPIC	69
10	IMAGE, COLOUR, LANGUAGE – THE LIMITATIONS OF SEMIOTIC SYSTEMS	71
10.1	THE LIMITATIONS OF COLOUR	71
10.2	THE VARIANCE OF MEANING OF IMAGES.....	73
10.3	THE LIMITATIONS OF WRITTEN LANGUAGE	78
11	ÜBERBLICKSLITERATUR /LITERATURE	80

1 EINLEITUNG

Dieser Teil des Leitfadens soll den im Frühjahr ausgefallenen Workshop zumindest in Teilen auffangen. Er widmet sich daher inhaltlich noch einmal der Multimodalität. Nachdem im zweiten Teil des Leitfadens aus dem Frühjahr sich der Funktionsweise multimodaler Artefakte oder Kommunikaten gewidmet und damit die Grundlagen gelegt hat, wendet sich dieser dritte Teil der multimodalen Praxis zu und erläutert an Beispielen unterschiedlicher Komplexität, wie Text und Bild zusammenarbeiten. Dabei kommt die Sprache nicht nur auf die direkte Interaktion von Text und Bild/Grafik und den entsprechenden Beziehungsgefügen, sondern auch auf benötigtes oder geliefertes Hintergrundwissen, das der Betrachter zum Verständnis braucht, zu sprechen. Die Inhalte dieses Leitfadens sollen es ermöglichen, die Komposition von Kommunikaten zu verstehen und so Interpretationshilfen für Daten liefern.

Kapitel 2 widmet sich dabei den Aneignungswegen von Kommunikaten und den zugrunde liegenden kognitiven Mechanismen. Kapitel 3 schlägt dann den Bogen zur Sehschulung und damit zum Zusammenspiel von kognitiven Prozessen und Konventionalisierung. Hierauf aufbauend wird dann in Kapitel 4 die Leistungsfähigkeit multimodaler Daten betrachtet. Dabei werden auch solche Aspekte angesprochen, die keinen Werbecharakter haben, um die Reichweite und Varianz von Kommunikaten deutlich zu machen. Das abschließende Kapitel 5 beschäftigt sich mit den Grenzen einzelner Modalitäten und zeigt, was das für die Konstruktion von Kommunikaten bedeutet.

2 TEXT UND BILD

Die Analyse von Kommunikaten ist Teil der Bildlinguistik, die eine eigenständige Disziplin ist, welche an der Schnittstelle von Soziolinguistik, Textlinguistik und Medien- und Kommunikationswissenschaften angesiedelt werden kann. Mit den Themen der Bildlinguistik befasst sich die Linguistic Landscape Forschung nur am Rande.

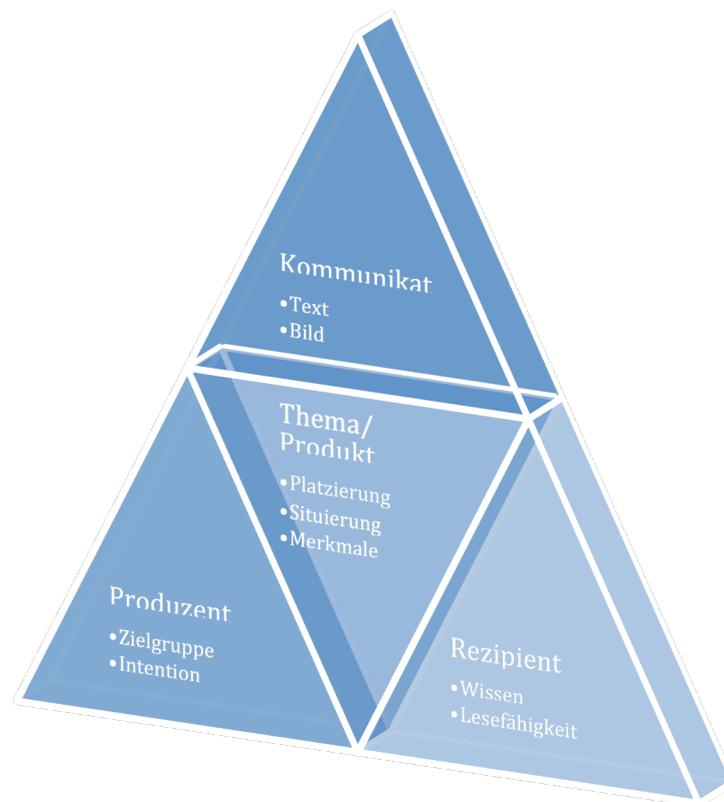
Eine wichtige Diskussion betrifft die Definition des Kernbegriffs „Text“. Im herkömmlichen Sinne handelt es sich dabei um ein rein sprachliches Etwas, das bestimmte inhaltliche und formale Anforderungen erfüllen muss, also Kohärenz und Kohäsion betreffen. Im Bereich der Bildlinguistik und der Multimodalitätsforschung ist aber der Beitrag weiterer Modalitäten zu Kohärenz und Kohäsion komplexer Kommunikate zu befragen, deren Aussage die Umfang eines Satzes übersteigt und die evtl. erzählerischen oder darstellenden Charakter haben.

Dabei ist auf der Ebene der nicht-sprachlichen Zeichen zu trennen zwischen Code und Mode/Modalität, da beide gelegentlich identisch gebraucht werden. So ist Farbe eine Modalität. Es gibt aber Farbcodes. Diese beziehen sich auf bestimmte konventionalisierte Bedeutungen bestimmter Farben: schwarz als Trauerfarbe, rot als Farbe der Liebe, weiß als Farbe der Unschuld etc., womit den Farben dann Symbolcharakter zukommt, wie es z.B. auch für Blumen bekannt ist: rote Rosen als Symbol für Liebe, Gladiolen für Stolz etc.

Sowohl als eigenständiger Untersuchungsgegenstand als auch als Teil der Analyse von Kommunikaten sind folgende Fragen zu berücksichtigen:

- Wie tragen welche Zeichen/modes zur Konstruktion des Kommunikats bei?
- Welche Rolle spielen
 - Produzent
 - Welche Zielgruppe hat er im Blick?
 - Was ist seine Intention?
 - Rezipient
 - Welche Fähigkeiten bringt er mit?
 - Was erregt seine Aufmerksamkeit?
 - Produkt
 - Situation, in der der Rezipient auf Produkt trifft
 - Expliztheit des Aggregats

Der zentrale Aspekt ist hier das abrufbare Weltwissen: Der Produzent verlässt sich darauf, dass der Rezipient über bestimmte Wissensbestände verfügt. Das Produkt verweist mehr oder weniger direkt auf diese Wissensbestände. Der Rezipient wiederum benötigt diese Wissensbestände zum korrekten (vom Produzenten intendierte) Verständnis des Aggregats.



2.1 Aneignung Von Kommunikaten:

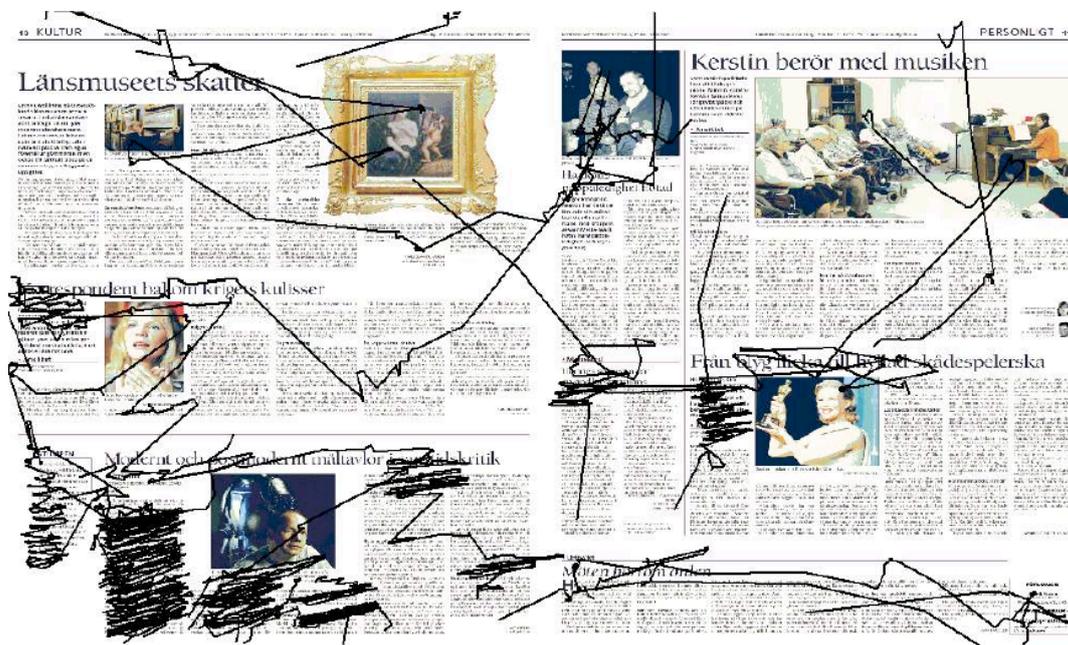
Je nach Kontext der Wahrnehmung des Aggregats können, wie erwähnt, Themen- oder Ortsgebundenheit sowie persönliche (idiosynkratische) Erfahrungen und Interessenslagen eine wichtige Rolle spielen.

Für Kommunikate gelten darüber hinaus zwei kognitiv basierte Regeln, die auf die Konstruktion Einfluss haben:

1. Bilder werden immer angeschaut
2. Bilder mit Gesichtern werden intensiver und länger angeschaut als Bilder ohne Gesichter

Aus der Summe der genannten Faktoren ergeben sich unterschiedliche Möglichkeiten, nach denen ein Betrachter in einem Kommunikat Einheiten bilden kann. Die Untersuchung der Bildung von Einheiten geschieht in unterschiedlichen Formen mittels Eytracking.

Das erste Beispiel aus Holmqvist & Wartenberg (2005) zeigt, wie eine Doppelseite einer Zeitung wahrgenommen wird. Zu erkennen ist, dass alle Bildelemente zumindest kurz betrachtet werden. Ebenso werden Überschriften stärker als andere Textteile betrachtet. Dies kann u.a. damit begründet werden, dass Überschriften einerseits aufgrund ihrer Größe auffälliger (salienter) sind als Fließtext und daher mehr Aufmerksamkeit auf sich ziehen und andererseits gelerntes Wissen um die Funktion hervorgehobener Textteile in Zeitungen zum Tragen kommt.



Holmqvist & Wartenberg (2005: 6)

Figure 4: Reading pattern over a broadsheet spread recorded as absolute data.

Das zweite Beispiel zeigt ein sogenanntes Simultanbild zu einer Geschichte aus einem schwedischen Kinderbuch. Simultanbilder sind Bilder, in denen Aktivitäten, die zeitlich aufeinander folgen in einem Bild dargestellt werden. Hier sind es die Vorbereitung eines Beetes zur Aussaat und die Saat selbst (Boerris & Holsanova 2012; alle folgenden Abbildungen aus Bateman 2014:247 nach Boerris & Holsanova 2012).

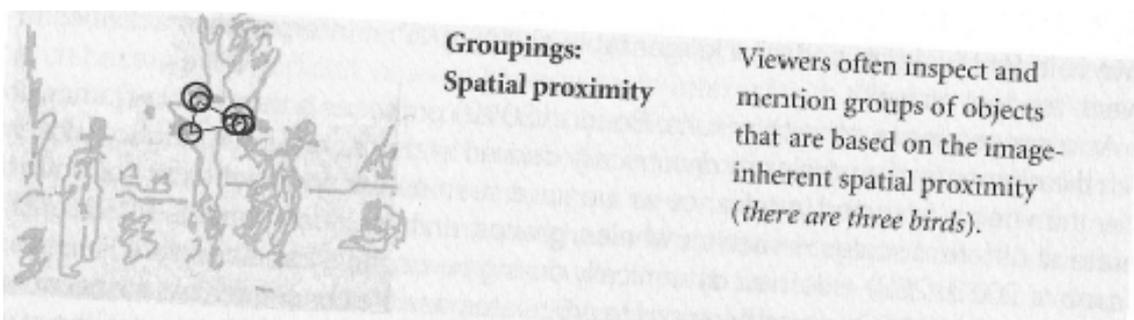
Dieses Bild steht für Untersuchungen, in denen der Frage nachgegangen wird, auf welcher Basis Aggregate gebildet, also Objekte in einen Zusammenhang gestellt werden. Dabei lassen sich einige grundlegende Mechanismen feststellen, die auf das kognitive Verfahren der Kategorisierung zuzuführen sind.



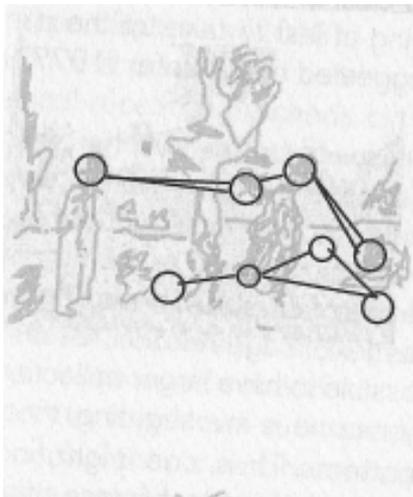
Figure 3 Complex picture: the motif comes from a children's book by Sven Nordqvist (1990). Reproduced by kind permission from Sven Nordqvist.

Die Wahrnehmungsreihenfolge und -intensität bestimmt durch:

Räumliche Nähe



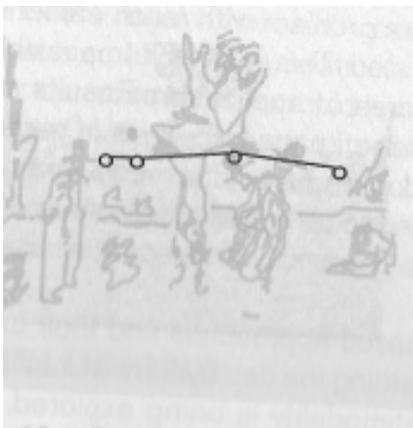
Kategorielle Ähnlichkeit (also alle Personen, alle Tiere, alle Bäume etc.)



**Groupings:
Categorical
similarity**

However, viewers also perceive and describe clusters of objects that are *not* spatially close. These clusters consist of groupings of multiple similar objects distributed over the whole scene (*four cats, four versions of Pettson*). This type of cluster is based on categorical similarity.

Mehrfach präsente Objekte

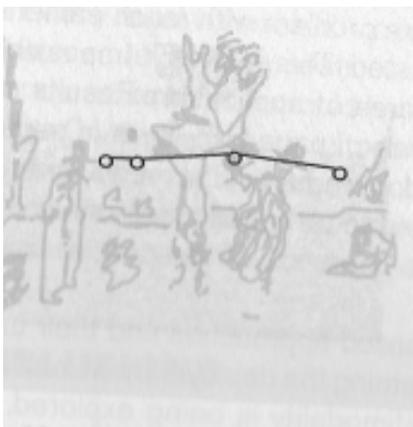


**Grouping:
composition**

activities.

Another type of cluster perceived as a meaningful unit in the scene is *hills on the horizon*. The viewer's eyes follow the horizontal line, filling in links between objects. This cluster seems to be a compositionally guided grouping.

Komposition entlang von Linien (Horizont)

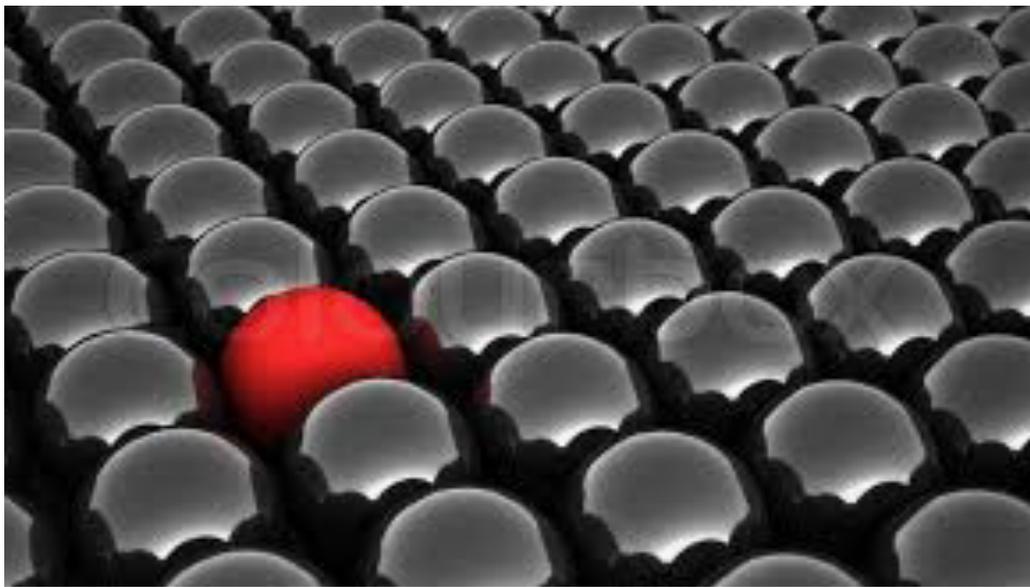


**Grouping:
composition**

activities.

Another type of cluster perceived as a meaningful unit in the scene is *hills on the horizon*. The viewer's eyes follow the horizontal line, filling in links between objects. This cluster seems to be a compositionally guided grouping.

Diese Verfahren haben eine kognitive und evolutionäre Grundlage, die sich auf die visuelle Aufmerksamkeit auswirkt und dazu führt, dass der Blick schnell und zuerst auf interessante Objekte gelenkt wird. Dies hat vermutlich evolutionäre Gründe wie z.B. die schnelle Entdeckung von Beute, Feinden oder Gefahren. Dabei können Stimuli in bestimmten Kontexten besonders auffällig (salient) sein. In den beiden nachfolgenden Bildern sind die rote Kugel bzw. die Krokusblüte nur durch ihre Umgebung salient. Die rote Kugel würde unter anderen roten Kugeln ebenso wenig auffallen, wie die lilafarbenen Krokusblüten und anderen gleichfarbigen Krokusblüten.



Boeris & Holsanova 2012



<http://bilder.4ever.eu/zeichentrick/digital-art/rote-kugeln-183874>



Boerris & Holsanova 2012



canstockphoto.at

2.2 Unterscheidung von Wahrnehmungsarten

Grundsätzlich können zwei unterschiedliche Wahrnehmungsarten unterschieden werden.

1. **Bottom-up:** Dieses Verfahren ist salienzbasierend. Die Wahrnehmung wird hier durch Auffälligkeiten ausgelöst und geleitet (vgl. die beiden obigen Beispiele).
2. **Top-down:** Dieses Verfahren ist bewusst und aufgaben-/interessensspezifisch. Das heißt, dass hier der Betrachter gezielt nach Informationen sucht und die Wahrnehmung von Objekten von ihrer Relevanz hinsichtlich der Aufgabe bzw. des Interesses abhängig ist.

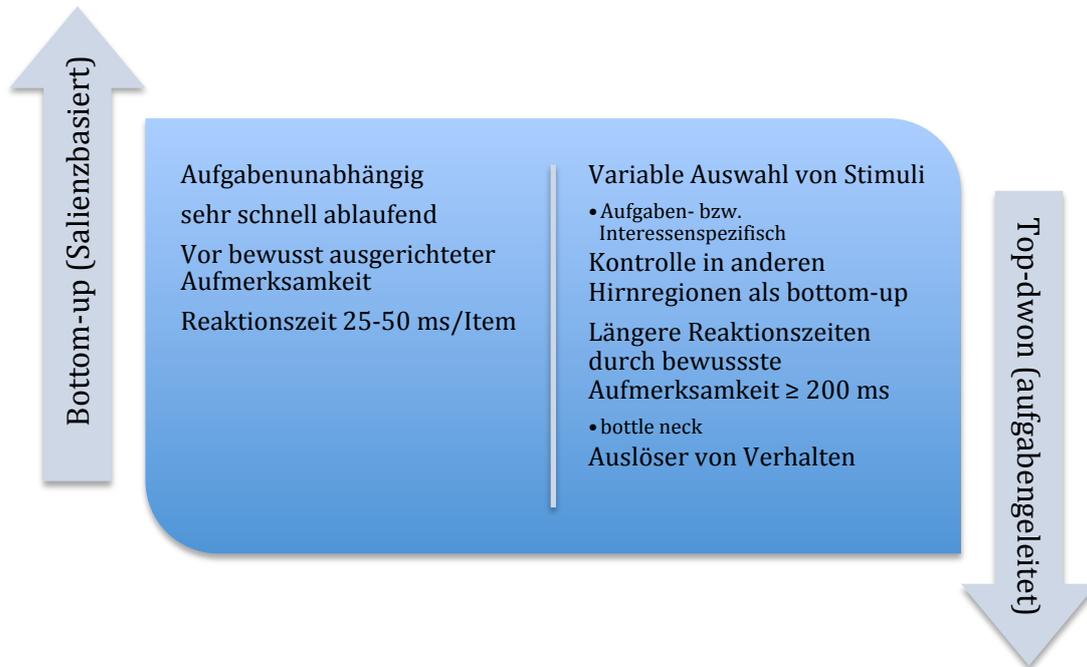
Die beiden Wahrnehmungsarten können kurz wie folgt charakterisiert werden:

Salienzbasierende Wahrnehmung

- unabhängig von Art oder Inhalt einer gegebenen Aufgabe
- läuft sehr schnell ab
- vornehmlich bottom-up gesteuert
- Verarbeitung vermutlich vor bewusst ausgerichteter Aufmerksamkeit
- Die Reaktionszeit liegt bei 25-50 ms/salientem Item

Top-down als Form der Aufmerksamkeit ist bewusster und mächtiger

- variables Auswahlkriterium für Stimuli
 - Aufgaben- und Interessensspezifisch
- Wird vermutlich in anderen Hirnregionen kontrolliert
- Die bewusste Aufmerksamkeit bedingt längere Reaktionszeiten: ≥ 200 ms
- Verursacht einen bottle neck
 - Daher keine vollständige gleichzeitige Verarbeitung der Datenflut
 - Statt dessen serielle Verarbeitung
 - Auflösen einer komplexen visuellen Szene in eine schnelle Abfolge einfacherer, lokal begrenzter Analysen
 - Ist Auslöser von Verhalten
 - Ist eng verbunden mit Erkennen, Planung und Bewegungskontrolle



Beide „Techniken“ werden bei allen visuell wahrnehmbaren Objekten angewandt:

- Nur Bilder
- Nur Text (Font, Schriftgröße, fett, unterstrichen, kursiv, Farbe etc.)
- Kombinationen aus Text und Bild

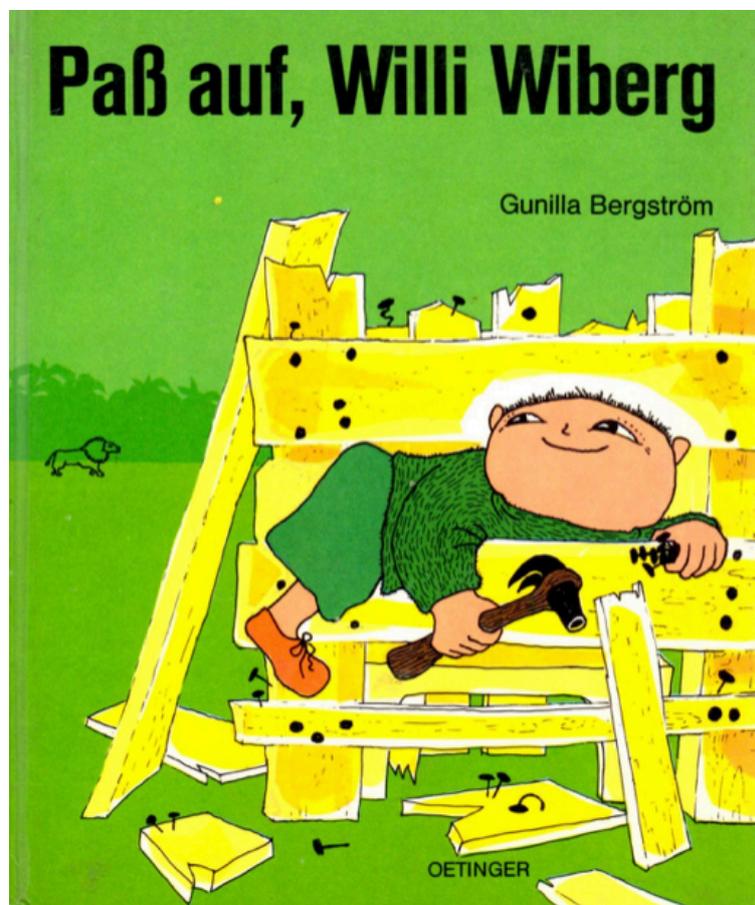
2.3 Text-Bild-Kombinationen

Sind die kognitiven Mechanismen bekannt, die die visuelle Wahrnehmung steuern, können Kommunikate bewusst komponiert werden. Dabei kommen Bildern/Grafiken bestimmte Funktionen zu. Die korrekte Interpretation von Bildern in Kommunikaten und damit die multimodale Lesefähigkeit ist eine Kompetenz, die von Kindheit an gelernt wird und die eine Kernkompetenz im Umgang mit Medien darstellt. Um deutlich zu machen, wie stark die visuelle Wahrnehmung von klein auf geprägt und geschult wird, werden die unterschiedlichen Funktionen anhand des Covers einer Geschichte aus der schwedischen Kinderbuchserie Willi Wiberg erklärt.

Bilder leisten mehr als nur den Text zu illustrieren:

- Bauen die Szene auf
 - Zeigen den Protagonisten bei einer für das Buch typischen Handlung
- Definieren/entwickeln Charaktere

- Willi Wiberg tut schon mal Dinge, die man nicht tun sollte – Willi ist ein Vorschulkind und hantiert mit Hammer und Nägeln. Das Lächeln zeigt, dass er offensichtlich Spaß an verbotenen Dingen hat
- Weiten die Handlung aus/entwickeln die Handlung
 - Willi Wiberg baut etwas, es geht als um Aktivitäten, die für Kinder gefährlich sein können, ihnen aber Spaß machen.
- Perspektivenwechsel
 - Der Titel ist die Perspektive der Eltern. Das Bild und das Lächeln auf dem Gesicht von Willi weisen auf eine andere Bewertung hin.
- Beitrag zur Textkohärenz
 - Text und Bild stehen in engem Zusammenhang
- Verstärkung der Textaussage
 - Hammer und Nägel können gefährlich sein, beim Umgang mit ihnen muss man vorsichtig sein. Das passt zum Titel.



Dabei tragen nicht alle grafischen/bildlichen Elemente unmittelbar zur Informationsstruktur bei. Es werden unterschieden:

- Dekorative Funktion
 - Einladung zum Kinderfest → Luftballons
 - Geringe bis keine Beziehung zum Text
- Repräsentative Bilder
 - Geben den ganzen Text oder Teile akkurat wieder
 - Die häufigste Form
- Organisatorische Bilder
 - Strukturierung der Textinformation
 - Arbeitsschritte
 - Karten
- Interpretative Bilder
 - Veranschaulichung komplexer Systeme oder Sachverhalte
 - Blutkreislauf
 - Technische Zeichnung
- Transformative Bilder
 - Häufig mnemonische Funktion
 - Bildliche Darstellung abstrakter Sachverhalte und Zusammenhänge
 - Grafische „Eselsbrücken“
 - Z.B. Darstellung des Ortsnamens *Eselsburg* durch das Bild eines Esels und einer Burg



© Bilder: Wikipedia

3 SEHSCHULUNG

Unter Sehschulung versteht man die gelernte profilierte Wahrnehmung und Interpretation außersprachlicher, visuell vermittelter Information. Diese beruht auf gesellschaftlicher Konvention und kann Moden oder historische Zustände widerspiegeln. Dabei kann z.B. Wertewandel zu Fehlinterpretationen oder Konflikten führen. Gleiches gilt beim Zusammentreffen unterschiedlicher kultureller/gesellschaftlicher Gruppen.

- Vgl. die unterschiedliche kulturelle Funktion von weiß
 - Europa: Freude, Reinheit
 - Teile Asiens: Trauerfarbe

Diese Konventionen wirken so stark, dass die Werbung internationaler Unternehmen immer lokalisiert wird. Lokalisierung meint die kulturspezifische Anpassung von Werbeinhalten in Text und Bild. In Folge der Globalisierung und der damit verbundenen Herausbildung einer allgemeinen, meist westlich dominierten Wissensbasis, ist es nicht notwendig für jedes Zielgebiet eine eigene Werbung zu entwickeln, sondern Kampagnen nur in den wesentlichen kritischen Punkten anzupassen.

Ein einfaches Beispiel ist die Aufladung von Fonts mit bestimmten Assoziationen:

- serifenbetonte Schriften stehen für Authentizität, Bodenständigkeit etc.
- Schriften ohne Serifen gelten als rational aber auch als edel

Diese Zuschreibungen beeinflussen die Fontwahl für Produkte:

Die beiden nachstehenden Labels wollen mit der Fontwahl zwei unterschiedliche Dinge ausdrücken. Während Captain Morgan mit der serifenbetonten Schrift den idealisierten unabhängigen Charakter von Piraten und deren Leben assoziieren will, geht es dem Whiskey-Label um die Betonung von Tradition. Diese Assoziationen können sowohl auf die Produktionsweisen und das Image der Labels als auch auf die Kunden bezogen werden. Sie sind daher sowohl für die Fremdwahrnehmung bzw. das Image der Marke als auch für die Selbstwahrnehmung der Kunden wirksam. Diese doppelte Leistung ist immer Gegenstand von Markenimages, die nicht nur ein bestimmtes Bild von sich selbst produzieren wollen, sondern darauf abzielen, dass sich die Zielgruppe mit diesem Bild identifiziert und die Assoziationen somit unbewusst auf sich selbst und damit das Selbstbild überträgt.



Die beiden folgenden Beispiele zeigen, wie Fonts im hochpreisigen Niveau eingesetzt werden. Das bekannte Chanel-Logo verwendet eine klare einfache Schrift ohne Serifen. Damit wird eine bestimmte Perspektive auf Luxus aufgegriffen, nämlich die Gleichung, dass „edel/elegant“ gleichgesetzt wird mit einer geradlinigen, klaren und auf den ersten Blick einfachen Erscheinung, deren Raffinesse im Detail liegt und nur für den Eingeweihten zu erkennen ist.



Das Label Bulgari geht einen etwas anderen Weg. Es verwendet einen Font mit leichten Serifen. Damit will es aber nicht die oben genannten Assoziationen wecken. Dass hier eine Distinktion möglich ist, liegt daran, dass der von Bulgari gewählte Font dem klassischen lateinischen Alphabet entspricht und an antike römische Inschriften erinnert. Zentral ist hier die Form „V“ für „U“. Damit wird ein gewisser Anspruch formuliert, da der Rezipient in der Lage sein muss, diese Form aufzulösen. Gleichzeitig

kann der Verweis auf das antike Rom nicht nur mit dem Firmensitz in der italienischen Hauptstadt, sondern auch mit einer bestimmten Lebensart assoziiert werden, da das Leben der höheren Kreise des römischen Reiches im Wissen der Allgemeinheit stärker verankert sind, als das Wissen um die Lebensumstände der einfacheren Römer.

Ähnlich, aber mit unterschiedlichen Fonts, arbeitet das Logo des französischen Luxuslabels Hermès. Der Firmenname selbst ist in einer serifenbetonten Schrift gehalten. Die Rustikalität, die damit zum Ausdruck kommt, verweist auf die Wurzeln des Unternehmens im Sattlerhandwerk. Diese Verbindung wird zusätzlich über die Grafik im Logo aufgegriffen. Diese soll also nicht nur historisch auf eine wohlhabende Kundenschicht verweisen, sondern auf den ursprünglichen Tätigkeitsbereich. Die über die Grafik geleistete Assoziation zu wohlhabenden Gesellschaftsschichten (nur diese konnten sich die dargestellte Kutsche und Kleidung leisten) wird über den serifenfreien Font des Ortsnamens aufgenommen. Paris wird nicht nur mit Reichtum, sondern auch mit Eleganz und Mode assoziiert. Hierfür sollen mit dem Font die oben für serifenfreie Fonts genannten Assoziationen aktiviert werden.

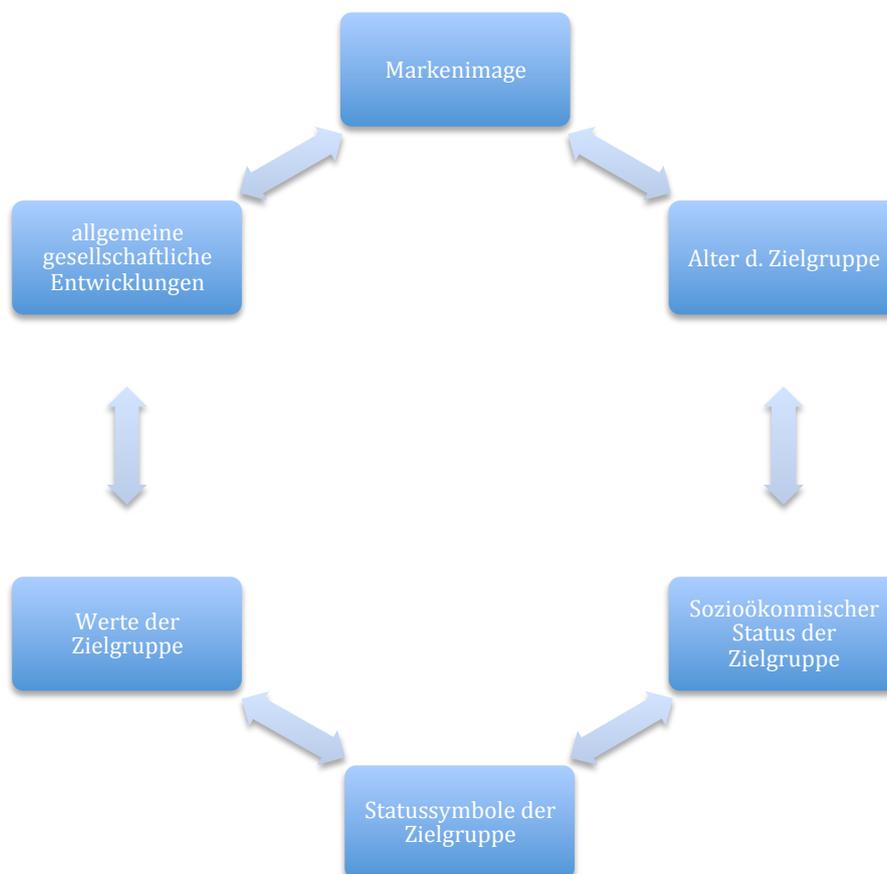
Es ergibt sich also in der Summe ein komplexes Logo, in dem die Fonts für bestimmte Werte stehen und die Grafik, obwohl im Kopf des Labels zu finden beide in sich vereinigt.



Wichtig ist, immer zu bedenken, dass die Interpretation von Schriften, Farben etc., wie oben gezeigt, stets auf der Basis der gelernten Assoziationen erfolgt. Dies geschieht unbewusst, automatisiert, da diese Assoziationen zum überindividuellen Wissensbestand gehören. Da es gleichzeitig ein Bewusstsein für dieses Wissen gibt,

erfolgt auf der Produzentenseite die Wahl einer Schrift, einer Farbe, eines Logos etc. immer im Wissen um diese gesellschaftlichen Konventionen um gezielt ein Markenimage aufzubauen und eine hohe Identifikation der Kunden mit der Marke und damit Kundenbindung zu erreichen. Markenimage und das Profil der Zielgruppe bedingen sich also gegenseitig. Bei der Neueinführung einer Marke bedeutet dies, dass zunächst eine Zielgruppe bestimmt werden muss, an deren Werten sich die Marke dann in der Außendarstellung orientiert. Im weiteren Verlauf muss eine Marke die Zielgruppe genau beobachten und bei Änderungen entsprechend reagieren.

Will eine Marke mit einer jungen Zielgruppe mit dieser „altern“ müssen sich Werbekampagnen etc. daran orientieren. Werden bestimmte Statussymbole obsolet muss ich die gesamte Außendarstellung damit auseinandersetzen, da die vom Produzenten gezielt aktivierten Assoziationen führen zu Kategorisierungen, die sowohl die Marke als auch die Kunden bestimmten Segmenten zuordnen:



Das nachfolgende Beispiel des Cafés der Firma Riquet aus Leipzig zeigt, wie die Außendarstellung sowohl die Firma repräsentiert als auch auf zur Entstehungszeit aktuelle Strömungen aufnahm und die Zielgruppe ansprach. Auf der Firmenwebsite wird die heute aus der Zeit gefallene, und kontextlos eher als „kitschig“ oder aber kolonialistisch bewertete Architektur zeitlich verortet und erklärt und somit zu einer historischen Aussage.

Herzlich Willkommen im Riquet-Haus

Das 1908/09 im Auftrag der Firma Riquet & Co vom Architekten Paul Lange errichtete Geschäftshaus gehörte seinerzeit mit zu den modernsten und zweifellos originellsten Neubauten Leipzigs.

Die bis ins Jahr 1745 zurückreichende Handelstradition der Firma Riquet mit Ostasien und dem Orient setzte der Architekt auf einfallsreiche Weise um.

Das doppelstöckig geschweifte Dachtürmchen entlehnte er der klassischen chinesischen Baukunst. Brüstungen, Pfeiler, Hauptgesims und Rundgiebel sind mit wunderschönen farbigen Mosaiken in Jugendstilmanier geschmückt. Zwei, die Eingangstür zum Cafehaus flankierende, kupfergetriebene Elefantenköpfe stellen das Markenzeichen der Firma Riquet dar.

<http://www.riquethaus.de/kaffeehaus.php>



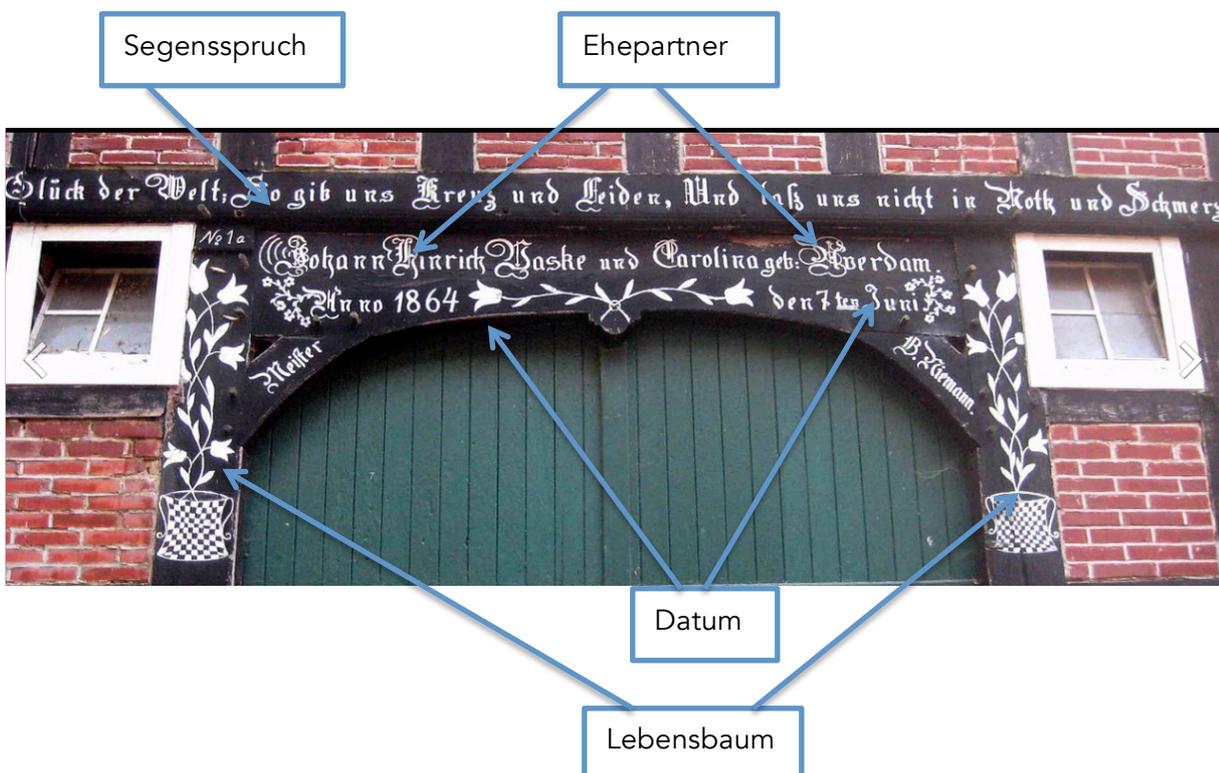
Dieses Beispiel zeigt auch, warum die Architektur in einer multimodal ausgerichteten Analyse des öffentlichen Raums unbedingt zu berücksichtigen ist. Die Einbettung der Architektur kann dabei unterschiedliche Bezugspunkte haben. Sie kann wie im Beispiel der Firma Riquet die Firmengeschichte widerspiegeln, aber auch auf allgemeine gesellschaftliche und/oder politische Ereignisse auf unterschiedlicher Ebene Bezug nehmen (Familiengeschichte, Ortsgeschichte, lokale Sagen, traditionelles Handwerk etc.).

4 BILDER, FARBEN, SPRACHE - DIE AUSSAGEKRAFT MULTIMODALER DATEN

Multimodalität bedeutet immer das gleichzeitige Auftreten von Informationen, in jeweils unterschiedlichen semiotischen Systemen, wobei alle Einzelinformationen zusammen die Gesamtaussage bilden. Dieser Ansatz erlaubt aber auch eine erweiterte Analyse von klassischen Linguistic oder Semiotic Landscape Daten und erweitert damit die Möglichkeiten, die Verwendung von Sprache und Diskursformen im öffentlichen Raum zu untersuchen. Dieses Kapitel widmet sich daher Beispielen multimodaler Praxis im öffentlichen Raum und zeigt an unterschiedlichen Beispielen, wie Multimodalität funktioniert und worauf bei der Analyse geachtet werden muss.

4.1 *Kommunikate als Spiegel sozialer Strukturen und Traditionen*

Das nachstehende Beispiel zeigt kulturelle Praxis als Teil des öffentlichen Raumes. In manchen Regionen Deutschlands war bzw. ist es üblich, nach der Eheschließung das Wohnhaus des Paares entsprechend zu schmücken. Das Haus mit einem religiös basiertem Segensspruch den Namen der Ehepartner, dem Jahr und einem Lebensbaum zu versehen, der symbolisch auch für die neue Familie steht, die wachsen und gedeihen soll



Ohne Kenntnis der zugehörigen Traditionen ist das Kommunikat aber nicht aufzulösen. So könnte vermutet werden, dass der Segensspruch in keinem zeitlichen Zusammenhang mit dem Rest steht, die Lebensbäume könnten als reiner funktionsloser, dekorativer (s.o.) Blumenschmuck gesehen werden. Die Datumangabe könnte als Bezugsdatum des Hauses und die Namen könnten als die der Bauherren interpretiert werden.

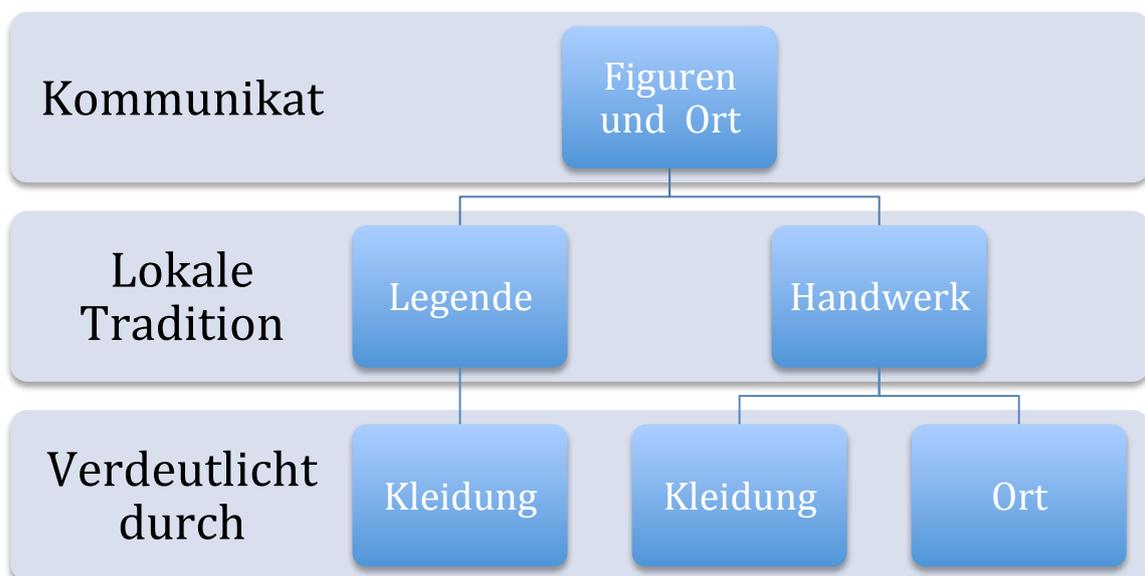
In diesem Beispiel werden aber Sprache und kulturell aufgeladene Symbole, zu der auch das Haus, hier als „zu Hause“, „Heim“ gehört zu einer Aussage vereint, in dem das Private öffentlich wird und somit in die geltenden gesellschaftlichen Normen und Erwartungshaltungen eingebunden wird. In der Linguistic Landscape eines entsprechenden Ortes hatten solche Strukturen nicht nur historischen Wert und sondern strukturierten den Ort, stellten öffentlich Beziehungen zu anderen her (Verwandtschaft) und können, sofern in ausreichender Zahl vorhanden, das soziale Netz des Ortes und seine Kontakte nach außen öffentlich darstellen.

4.2 Sprachfreie Präsentation lokaler Legenden und Traditionen

Das zweite Beispiel zeigt die Integration lokaler Sagen/Traditionen und Geschichte in den öffentlichen Raum. Die auf dem Bild gezeigte Abzweigung der Schäfflerstraße von der Weinstraße in München wird im Volksmund „Schäfflereck“ genannt.



Zum einen verweist die Position der Figuren an der Abzweigung zur Schäfflerstraße auf historische Strukturen. Straßennamen, die Gewerbe in sich tragen, verweisen häufig darauf, dass in der entsprechenden Straße früher dieses Gewerbe ansässig war. Entsprechend verweisen die Figuren am Eingang der Schäfflerstraße auf die Straße, in der in früheren Zeiten in München die Schäffler ansässig waren. Besonders ist aber die Tracht der Figuren, die nicht der traditionellen Arbeitskleidung entspricht, sondern der Tanzkleidung der Schäffler. Hiermit wird eine Stadtlegende aufgenommen, nach der die Schäffler Münchens 1517 während einer Pestepidemie erstmals öffentlich einen Tanz aufgeführt haben sollen, um die Menschen aus den Häusern zu locken und unterhalten. Der Schäfflertanz wurde seit dem in unregelmäßigen Abständen, seit 1760 aber alle sieben Jahre aufgeführt. Hinweise auf den Schäfflertanz aus dem Jahr 1517 finden sich auch an weiteren Stellen in der Münchner Innenstadt. Dass die historische Richtigkeit über die Entstehung des Schäfflertanzes nicht gesichert ist, spielt eine untergeordnete Rolle, da das ganze als Legende verstanden wird.



Die beiden Schöfflerfiguren sind sprachfrei und erhalten ihre Bedeutung und Berechtigung zum einen durch den Aufstellungsort, der auf historische Münchner Strukturen verweist und damit auch in umgekehrter Richtung eine Identifizierung der Figuren ermöglicht. Zum anderen verweisen die Figuren durch ihre spezielle Kleidung ein anderes, nicht mit dem Handwerk selbst verbundenes Ereignis aus der Münchner Stadtgeschichte. Damit erzählen die Figuren eine Geschichte ohne die Verwendung von Sprache.

4.3 Aggregate aus Text und Architektur

Das dritte Beispiel hat eine ganz andere Reichweite und gehört in den Komplex der historischen Monumente. Obwohl auch Beispiel 1 eine historische Dimension hat, wird es nicht als historisches Monument gezählt, da es keine „nationale“, politische Dimension hat.

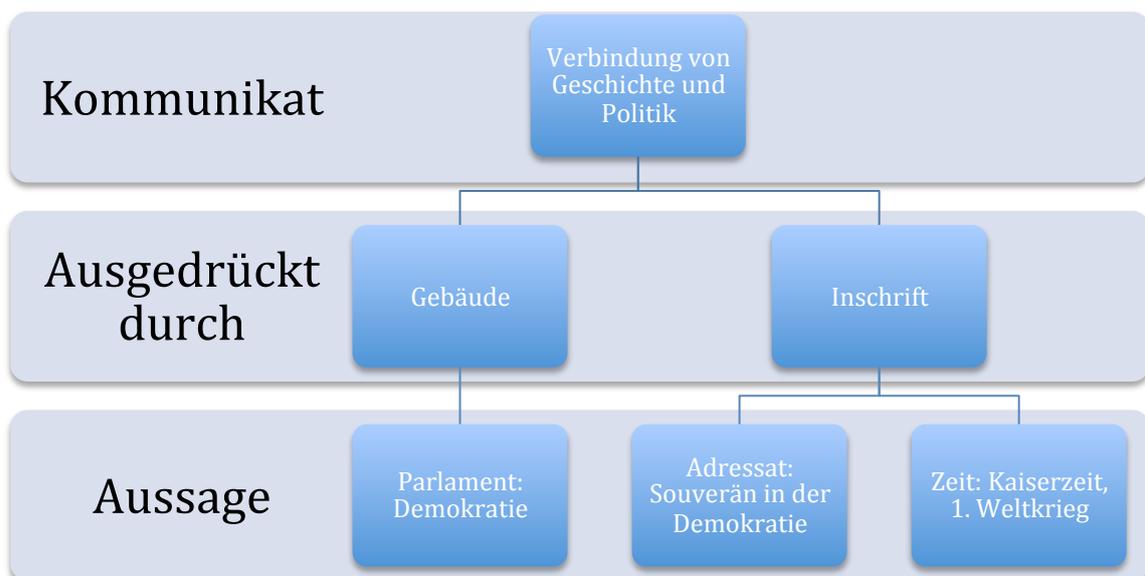


Bilder: https://de.wikipedia.org/wiki/Dem_deutschen_Volke

Die Inschrift findet sich über dem Westportal des Reichstagsgebäudes in Berlin. Damit ist der zentrale politische Ort einer Demokratie, nämlich das Parlament seinem Souverän gewidmet. Die 1916 angebrachte Inschrift ist durch ihren Ort einerseits zeitlos, hat aber auch eine eigene, zusätzliche historische Dimension, wenn bedacht wird, dass sie während des ersten Weltkrieges und noch zu Kaiserzeiten angebracht wurde.

Ihre zentrale Aussage erhält die Inschrift also nur im Zusammenspiel mit dem Gebäude bzw. dessen Funktion. An anderen Stellen wäre die Inschrift bestenfalls seltsam oder gar unverständlich. Unterstützt wird die Aussage durch die Beflaggung des Gebäudes, die auf die hoheitliche, nationale Bedeutung des Gebäudes verweist, so dass auch ohne Kenntnis um die Funktion des Gebäudes auf dessen bedeutenden öffentlichen Charakter geschlossen werden kann, da ansonsten Beflaggungen in Deutschland unüblich ist.

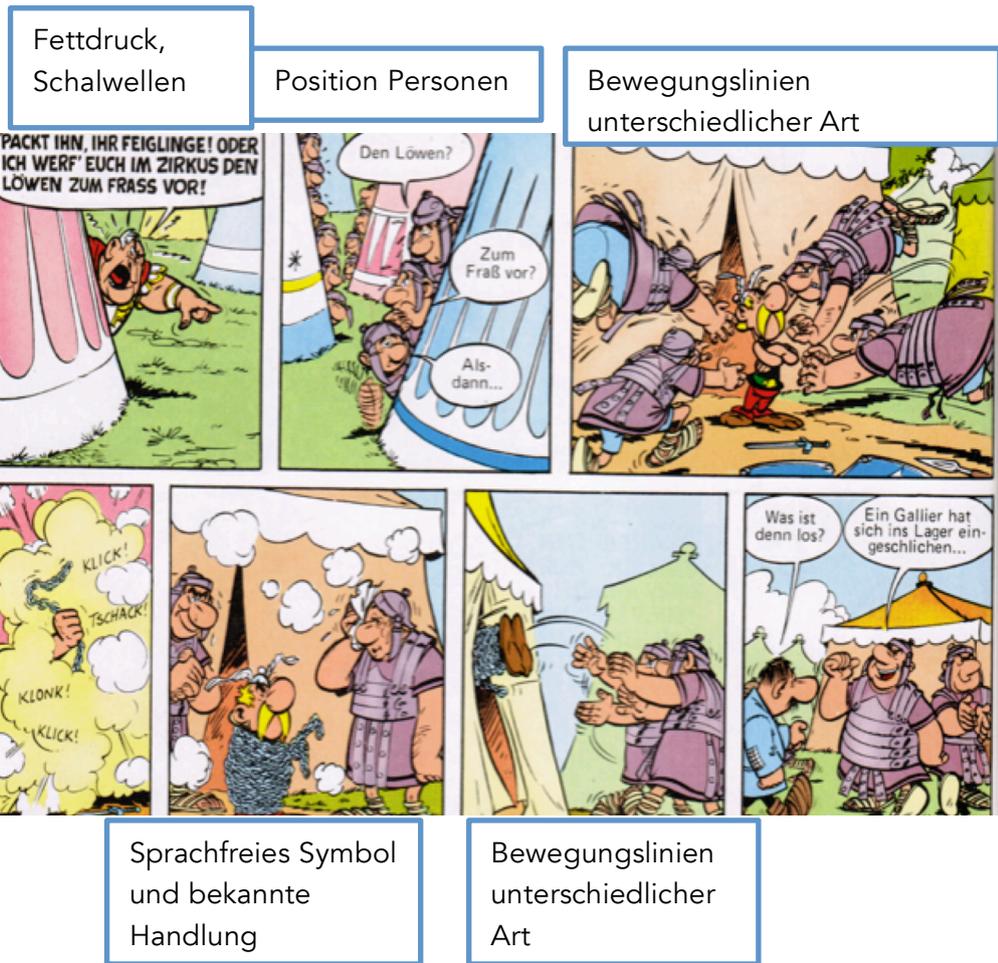
In diesem Beispiel interagieren also Schrift, Ort der Anbringung und zusätzliche Symbole miteinander und erklären sich in ihrer Präsenz gegenseitig. So kann aus Inschrift und Flaggen auf den Charakter des Gebäudes geschlossen werden, während dieser gleichzeitig Voraussetzung für die Präsenz des Textes und der Flagge ist.



4.4 Funktion und Bedeutung non- und paraverbaler Elemente

Dieses Beispiel stammt aus einem Comic und verdeutlicht damit, dass die Lesefähigkeit multimodaler Aggregate von Kindheit an gelernt wird. Dies bezieht sich sowohl auf die Interpretation nicht-sprachlicher Elemente als auch die Konstruktion von Bedeutung aus dem Zusammenspiel nicht-sprachlicher und sprachlicher Elemente.

Wie bedeutsam nicht sprachliche Elemente sind, zeigt sich, wenn man sich die Bilder ohne sie vorstellt.



<http://www.erzaehlenbilder.de/2015/01/06/zack-bummi-knuff-ganz-einfach-oder-doch-nicht/>

Fettdruck, Versalien =
Brüllen

Wird semiotisch durch das
Bild unterstützt

- Schallwellen für
Lautstärke
- Kleidung: Sprecher =
Offizier
- Ängstlich = Versteckt
sich hinter dem Zelt



Insgesamt wirken die einzelnen Modalitäten also wie folgt zusammen:

1. Zusammenspiel von Text und Typographie

Der Text ist ein Befehl. Dies wird einerseits durch die Formulierungen deutlich. Dass dieser Befehl nicht ruhig gesagt wird, sondern gebrüllt, wird durch den Fettdruck und die Versalien angezeigt. Diese Darstellung des Textes unterstützt die Formulierung. Typografische Aspekte haben also großen Anteil an der richtigen Interpretation der Aussage und verstärken bzw. unterstützen diese. Ziel ist die semiotisch gestützte Imitation der entsprechenden Sprachhandlung.

2. Nonverbale Charakterisierung des Sprechers

Dass die dargestellte Figur die Autorität besitzt Befehle zu erteilen, wird durch die Kleidung gezeigt, die die Figur als Offizier ausweist. Die Positionierung der Person stellt dann einen Widerspruch her, da die Figur offensichtlich Angst vor etwas hat, sich selbst in Sicherheit bringt, Ihre Leute aber genau dieser Gefahr aussetzt. So wird erneut eine Beziehung zwischen der Figur und dem Text hergestellt.

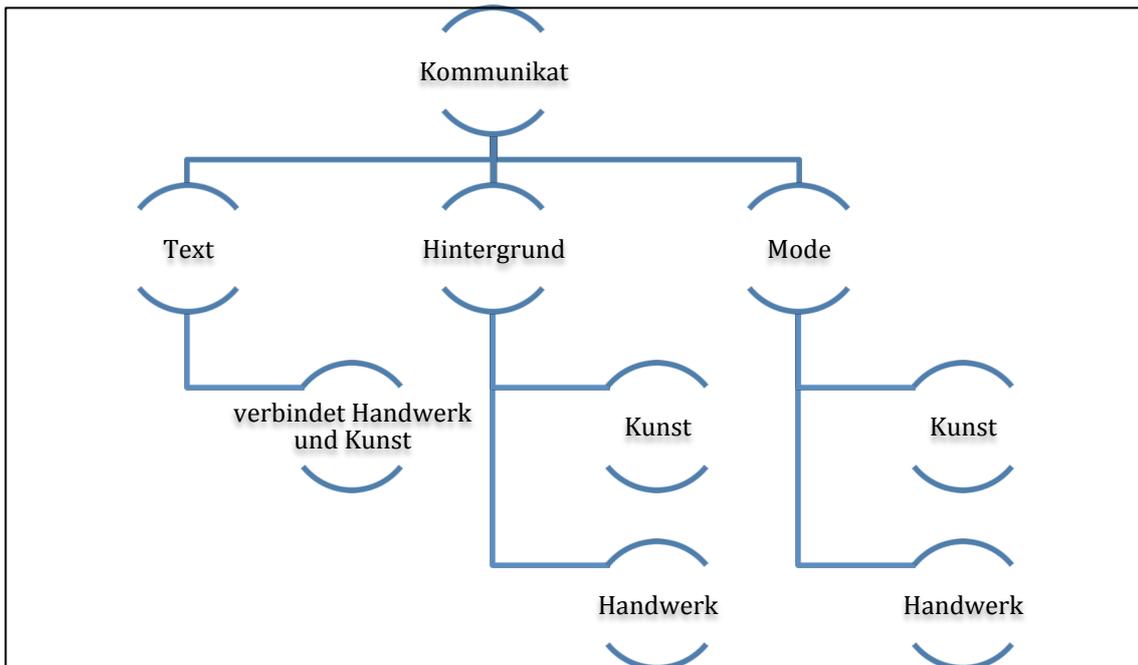
4.5 Zusammenspiel mehrerer Komponenten in komplexen Kommunikaten

Dieses Beispiel zeigt, wie komplexe Werbeaussagen im öffentlichen Raum auf mehreren Ebenen konstruiert werden können.

Der Text auf der Scheibe bezieht sich auf den Hintergrund, der nicht einfach ein grünes Ziermuster ist, sondern Kunst auf Basis von Lotusblättern. Der Text befasst sich inhaltliche auf Kunst, Handwerk und Natur und die handwerkliche Veredelung letzterer, die so zu Kunst wird. Damit referiert er auf den Hintergrund. Der Bezug wird durch das grafisch dargestellte Lotusblatt an der oberen rechten Ecke des Textes. Zwischen Text und Kunstwerk ist Kleidung positioniert, die, wenn man eine komplexe Aussage konstruieren will, auch als Kunst beschrieben werden kann, vgl. hierzu Deutsch *Handwerkskunst* → *Schneiderkunst*.

Der Text ist nur aus der Nähe gut lesbar, so dass bereits aus geringerer Entfernung eine komplexe Interpretation des Schaufensters nicht mehr möglich ist. Grundsätzlich wahrnehmbar ist dann nur der grüne Hintergrund mit der vor ihm ausgestellten Kleidung.



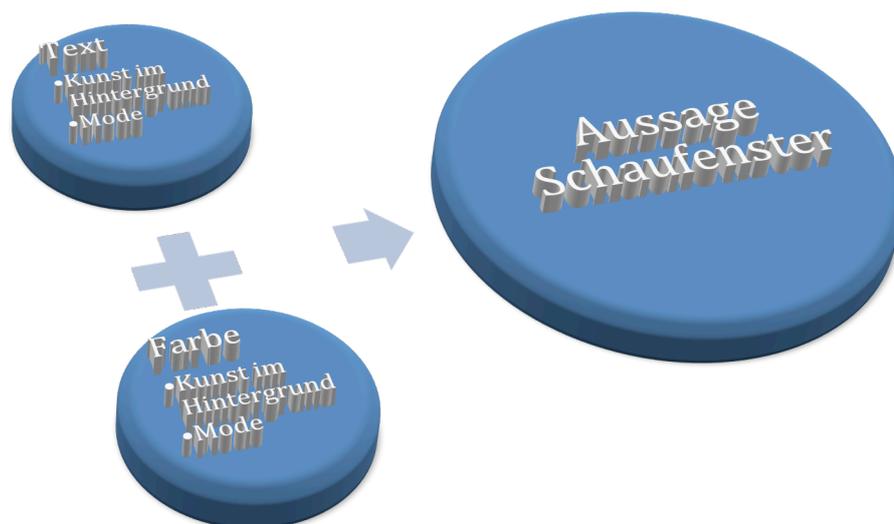


Die gleiche Komposition findet sich im benachbarten Schaufenster. Anders sind hier lediglich die Farbe des Kunstwerks, das nun golden und nicht grün ist und der Typ der gezeigten Kleidung. Auch hier ist für das Verständnis der komplexen Botschaft große Nähe zum Schaufenster nötig. Bei Abstand zum Schaufenster ist nur der Hintergrund mit der ausgestellten Kleidung wahrnehmungsprägend.



Bei beiden Schaufenstern wird mit einem Kontrast von Hintergrund- und Kleidungsfarbe gearbeitet, so dass sich die ausgestellten Kleidungsstücke gut vom Hintergrund abheben. Weiter ist der mögliche Zusammenhang zwischen Hintergrundfarbe und Kleidungsstil zu beachten. Im ersten Fenster ist ein grüner Hintergrund gewählt worden. Grün steht prototypisch für Natur. Dazu passt gezeigte junge Freizeitmode. Diese Assoziation wird verstärkt durch die haltestellenartige Konstruktion, vor der die Halbpuppen gruppiert sind und die möglicherweise für Bewegung, Ausflug o.ä. stehen kann. Einen Kontrast bildet das zweite Fenster. Die hier gezeigte seriöse Kleidung wird vor einem gedeckten, goldfarbenen Hintergrund gezeigt.

Die beiden Schaufenster verdeutlichen, wie über den Umweg eines zweiten Metiers, nämlich der Kunst, die gezeigten Produkte dadurch aufgewertet werden, dass sie in einen unmittelbaren Kontext mit diesem gestellt werden. Für Rezipienten, die den Text nicht wahrnehmen, bleibt immer noch das Zusammenspiel von Farbe und Kleidungsstil, welches im ersten Fenster noch durch ein zusätzliches Element gestützt wird.



4.6 Nonverbale Betonung eines Themas

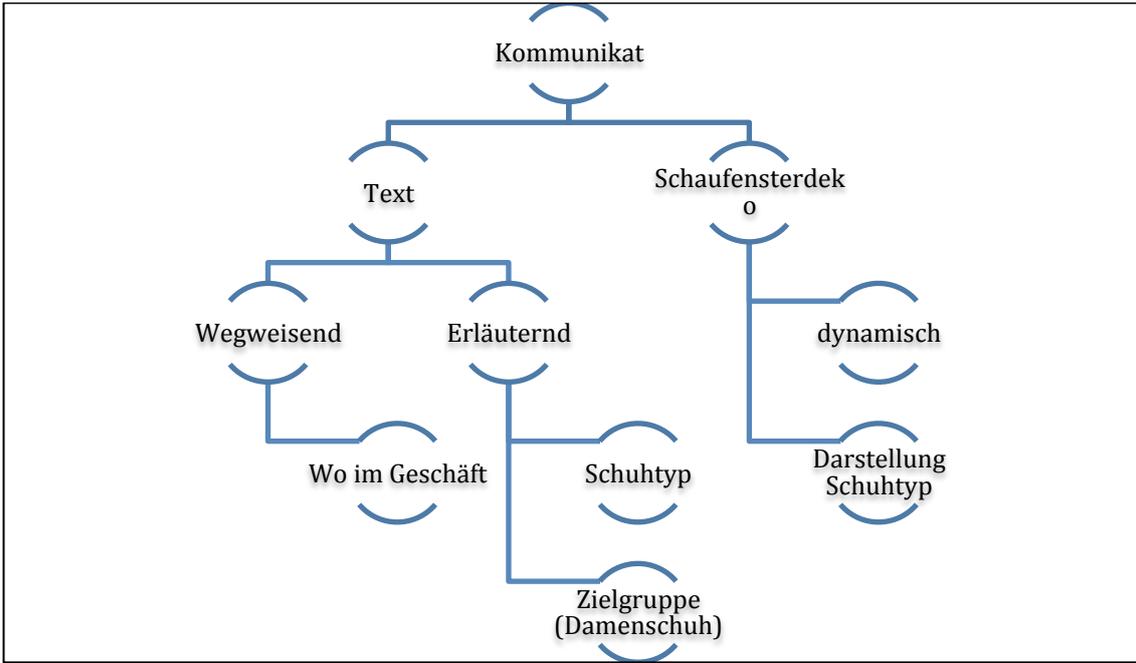
Das folgende Beispiel zeigt, wie das Thema eines Kommunikats nonverbal aufgegriffen und erweitert wird.



Der sprachliche Teil des Schaufensters ist sowohl erläuternd als auch wegweisend:

1. Den erläuternden Charakter hat der Teil „sneakers women“, weil die ausgestellten Produkte nicht eindeutig als Damenschuhe zu erkennen sind, sondern auch als geschlechtsneutral als Sneakers gesehen werden könnten.
2. Wegweisend ist die Geschossangabe, die darauf hinweist, wo im Geschäft die gezeigten Produkte erhältlich sind.

Die Anordnung der Produkte verweist ihrerseits auf das klassische Anwendungsgebiet dieses Schuhtyps, da die schräg gestellten Regale dynamischen Charakter haben und deuten damit Bewegung andeuten. Auch wenn Sneakers heute Teil bestimmter Modestile sind und nicht mehr ausschließlich im Rahmen sportlicher Betätigung getragen werden, sind sie entsprechend konnotiert.



5 BILDER, FARBEN, SPRACHE – DIE GRENZEN SEMIOTISCHER SYSTEME

Im vorhergehenden Kapitel ging es um die Aussagekraft von Kommunikation und die Interaktion der einzelnen Modalitäten. Die Konstruktion von Kommunikaten unterliegt dabei nicht nur den gelernten Konventionen, sondern auch den Grenzen der einzelnen semiotischen Systeme. Nicht jede Modalität kann alles (eindeutig) ausdrücken. Die nachfolgenden Beispiele sollen einen ersten Einblick in die Grenzen semiotischer Systeme geben und damit das Verständnis für die Komposition von Kommunikaten stärken.

Die Grenzen der einzelnen Modalitäten speisen sich aus zwei unterschiedlichen Komponenten.

1. Konventionalisiertes Wissen. Im nachfolgenden Beispiel ist dies z.B. die Bedeutung von Farbe.
2. Kognitive Verfahren.
 - a. Bilder werden holistisch also auf einmal vollständig wahrgenommen. Die Stelle, an der der Betrachter dann „tiefer“ in das Bild einsteigt, ist kontextabhängig und entweder salienzbasierend (bottom-up) oder aufgabenspezifisch (top-down (s.o.)). Bilder verarbeitungsbedingt besser memoriert als Sprache.
 - b. Schrift/Text wird immer sequentiell wahrgenommen und verarbeitet (also Wort für Wort). Dabei kann auch die Reihenfolge nicht beliebig gewählt werden, da sonst der Inhalt unverständlich wird.

5.1 Grenzen der Modalität Farbe

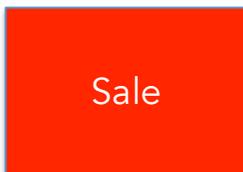
Farbe wurde im vorangegangenen Leitfaden als abhängige Modalität beschrieben, da sie alleine keinerlei Bedeutung besitzt, da ihre Erscheinung immer objektgebunden ist und in bestimmten Kontexten auftritt. Grundsätzlich kommt bei Farbe hinzu, dass sie sowohl eine konventionalisierte und damit in unterschiedlichen Gesellschaften unterschiedliche Bedeutung haben kann als auch eine aus der Natur abgeleitete Bedeutung haben kann.

Gesellschaftlich konventionalisiert ist der Bezug zur Liebe wohingegen Aufmerksamkeits- und Gefahrenhinweisfunktion auf dem Signalsystem der Natur beruhen. Demzufolge hat das kontextfreie rote Quadrat keinerlei spezifische

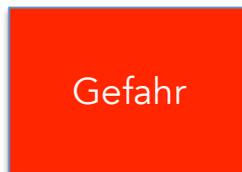
Bedeutung. Es kann für beide genannte Bedeutungen verwendet werden. Damit erhält die Farbe ihre endgültige, fallspezifische Bedeutung erst durch den hinzugefügten Text bzw. die hinzugefügten Grafiken.



Potentiell mehrere Bedeutungen:



1. Aufmerksamkeit



2. Warnfunktion



3. Emotion

Den Beispielen entsprechend kann diese Funktion von Farbe auch mit der Modalität Schrift verwendet werden, wenn diese entsprechend gefärbt ist. Hier sei das Beispiel aus dem vorhergehenden Teil des Leitfadens wiederholt, das darüberhinaus im Schaufenster die symbolische Bedeutung Liebe in einer dem dritten Beispiel oben analogen Form wiederholt.



5.2 Bedeutungsvarianz von Bildern

Ebenso wie Farben und andere abhängige Modalitäten sind Bilder in ihrer Bedeutung/Botschaft nicht eindeutig und können unterschiedlich eingesetzt werden. Das nachstehende Bild einer Blumenwiese kann so unterschiedliche Funktionen haben:



Im ersten Beispiel ist die Blumenwiese nur Hintergrund für das eigentliche Objekt, nämlich die Biene, die im Rettung bittet. So eingesetzt könnte das Bild Teil einer Umweltschutzkampagne sein. Die Blumenwiese ist dabei Teil der schützenswerten Natur, die auch von der Existenz von Bienen abhängt; dies wird aber nicht direkt thematisiert, so dass sie vordergründig nur als Habitat in Erscheinung tritt.



Auch in diesem zweiten Beispiel hat das Bild nur indirekt mit dem eigentlichen Thema zu tun, das sich mit dem Besitz eines Hauses auf dem Land beschäftigt. Allerdings verweist „Haus auf dem Land“ auch auf das Leben auf dem Land und den Umstand, dass zumindest in einer idealisierten Vorstellung Häuser dort Gärten haben und/oder die Natur unberührt ist. Daher ist das Bild eine Darstellung dieser idealisierten Vorstellung vom Leben auf dem Land und das Ende der Assoziationskette, die vom Text ausgeht.



Eine gegenteilige Bedeutung hat das Bild im nachstehenden Kontext, in dem es prototypisch für Auslöser der genannten Beschwerden steht.

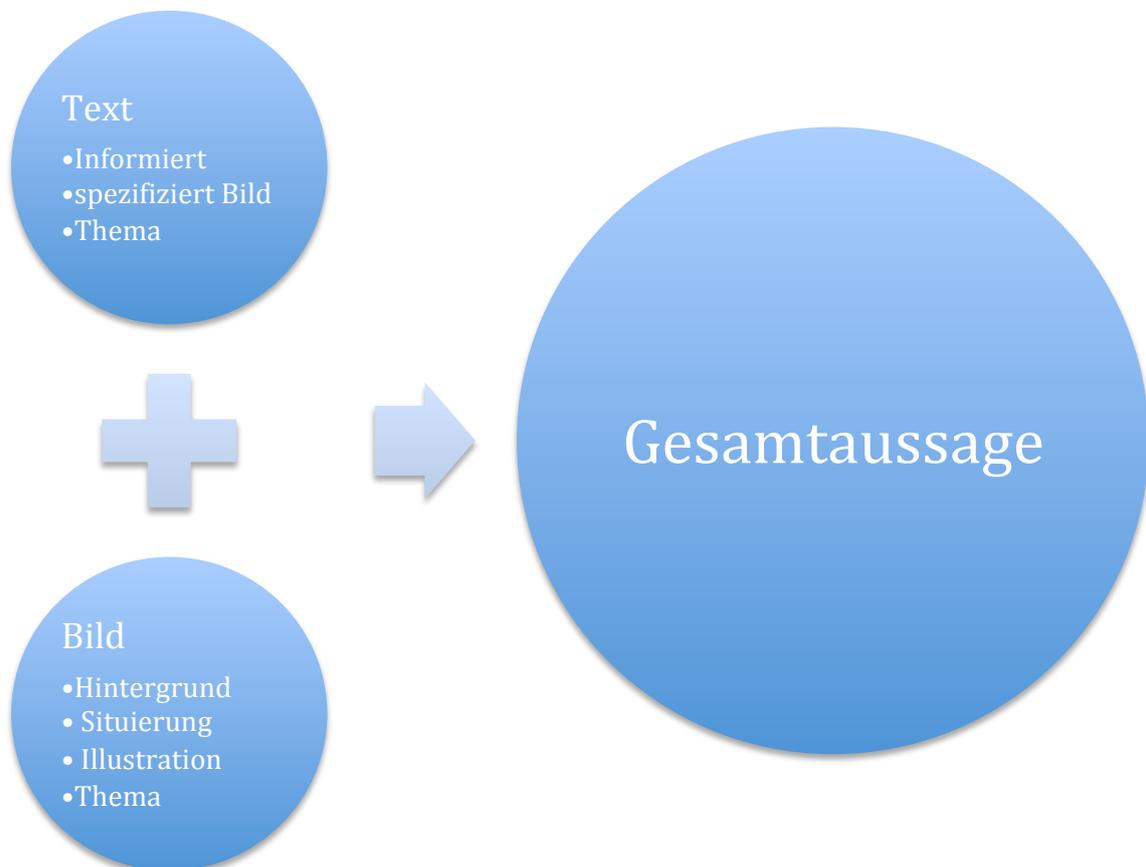


Gemein ist beiden Fällen, dass die Darstellung eine direkte Verbindung zu mindestens einem der Wörter im Text hat. Während im Beispiel 2 eine mehr oder minder direkte Verbindung zum Wort „Land“ gegeben ist, ist die Verbindung des Wortes „Sommer“ mit dem Bild als etwas stärker zu bewerten.

Wie bei den Beispielen zur Farbe Rot wird auch hier deutlich, dass Bilder, Grafiken etc. in ihrer konkreten Bedeutung oft kontextabhängig sind und je nach Komposition des Kommunikats eine unterschiedlich starke Rolle für die Bedeutung haben. Nicht nur das Beispiel mit der Biene wäre auch ohne das Bild bzw. mit einem anderen Bild denkbar. Dies zeigt, dass die Auswahl der Bilder/Grafiken gerade in der Werbung nicht nur von der Werbeaussage abhängt, sondern auch vom Markenimage und der Zielgruppe, die mit diesen Bildern angesprochen werden sollen.

Das Gesamt der Aussagen wäre auch rein textuell zu vermitteln. Dieser Prozess wäre aber länger, weil ein beschreibender Text gelesen und verarbeitet werden müsste. Gleichzeitig ist ein solcher Text weniger emotional als die bildlichen Darstellungen, die sofort Assoziationen wecken und dadurch viel unmittelbarer auch an die individuellen Erfahrungs- und Wissenswelt der Betrachter anschließen; vgl. hierzu die Aussage von Straßner (2002) wonach Bilder „schnelle Schüsse ins Gehirn [sind]“.

Auch das letzte Beispiel zeigt die emotionalisierende und kontextualisierende Funktion von Bildern. Die gezeigte Mode wird im nachstehenden Beispiel durch das Hintergrundbild und die Vespa im Vordergrund emotional aufgeladen und in den Kontext Reise/Italien etc. gestellt. Dabei wirken die Assoziationen, die mit einer Vespa verbunden sind. Wichtig ist in diesem Beispiel aber, dass diese Assoziationskette erst sprachlich ausgelöst werden muss.



Die Wahrnehmung beschleunigende Funktion von Abbildungen zeigt das nächste Beispiel:

Sie benötigen einen Hammer, zwei Nägel, dazu eine Säge und Leim.

Sie benötigen:

1 

2 

1 

1 

Der Vorteil des obigen Systems, das sich z.B. häufig in Bauanleitungen findet, ist zum einen, dass eine schnelle und gute Übersicht über das benötigte Material möglich ist und zum anderen Übersetzungen entfallen, da die Ziffern und Abbildungen universell verständlich sind.

Insgesamt zeigen die Beispiele die Polyvalenz von Bildern und grafischen Darstellungen. Zwar kürzen sie den Wahrnehmungsprozess ab, sind aber nur sehr eingeschränkt in der Lage eigenständig eine eindeutige Botschaft zu vermitteln. Hierzu benötigen Bilder immer eine Situierung, die eine Interpretations- bzw. Assoziationsrichtung vorgibt. Ausnahmen bilden hier nur solche Abbildungen, die eindeutig zu identifizieren sind und die in einer gegebenen Gesellschaft nur eine Interpretation zulassen. Aber selbst hier ist durch gesellschaftlichen Wandel nicht ausgeschlossen, dass sich Werte und Zuordnungen aufweichen. Als Beispiel mag hier der Umgang mit religiösen Symbolen dienen, die teilweise Modeaccessoires geworden sind, deren Verwendung nichts über die Religiosität des Trägers aussagt.

Dies gilt auch für den öffentlichen Raum mit seiner hochgradig konventionalisierten Themen- und Diskursstruktur. Sprachfreie Schaufenster sind für den Rezipienten nur verständlich, weil die Funktion von Schaufenstern und ihre Verbindung zu dem zugehörigen Geschäften bekannt ist. Entsprechend können die sprachlichen

Aufforderungen und Informationen wie „Hier gibt es“, „Hier können Sie kaufen“, „Treten Sie ein“ o.ä. entfallen, ohne dass es zu einem Informationsverlust kommt.

5.3 Die Grenzen verschrifteter Sprache

Die Ausführungen des obigen Kapitels haben gezeigt, dass Sprache in ihrer verschrifteten Form in Kommunikaten unerlässliche Interpretationshilfe für Bilder und Abbildungen ist.

Gleichwohl hat auch verschriftete Sprache in Kommunikaten ihre Grenzen. Die wichtigste ist wohl, dass alle parasprachlichen Funktionen nicht greifen und z.B. Emotionen, Ironie, Zweifel etc. nicht vermittelt werden können. Aus der Literatur ist bekannt, dass es hier immer „Lesehinweise“ braucht wie z.B:

„Lass gut sein!“, lachte, zischte er;

„Wir haben die Vase zerbrochen.“, gestanden Paul und Paula unter Tränen.

Für Kommunikate sind solche Ausführungen zu lang, da sie in der Regel darauf abzielen, von Passanten auch im Vorbeigehen nicht nur wahrgenommen, sondern später auch erinnert zu werden. Diesen Teil der Aufgabe übernehmen im Kommunikat dann vorzugsweise Bilder und grafische Darstellungen, die die entsprechenden Assoziationen und Emotionen aufrufen.



Gleiches gilt in Werbung auf für Wortspiele, wie das folgende Beispiel von Sixt zeigt. Unabhängig vom Übertragungsprozess, kann kein sprachliches Signal das Wortspiel andeuten. Anders als die o.g. Beispiele kann das Fehlen der Abbildung auch nicht zur ausführliche sprachliche Beschreibung ausgeglichen werden.



The advertisement features a gold Aston Martin Vanquish Coupé on a yellow background. The Sixt logo is in the top right corner. Below the car, the headline reads 'Englischer Rasen.' followed by the rental offer: '(Jetzt mieten: den Aston Martin Vanquish Coupé für € 189,-/Tag* inkl. 150 km unter sixt.de)'. A small disclaimer at the bottom right states: '* Bei einer Anmietung von 7 Tagen, zzgl. 23,9% Standortzuschlag bei Anmietung an Flughäfen, je nach Verfügbarkeit an ausgewählten Stationen.'

Die Orthographie des Deutschen ist zunächst korrekt, so dass das Wort „Rasen“ eindeutig zugeordnet werden kann, da „englischer Rasen“ eine Kollokation ist, die einen bestimmten Rasentyp beschreibt. Erst in Kombination mit der Abbildung des Autos, wird aus dem eine Grünfläche beschreibendem Nomen ein Verb, das ein Fahrverhalten beschreibt, dass mit dem gezeigten Autotyp assoziiert werden kann. Dabei ist es unerheblich, dass bei der Nationalitätsbezeichnung eigentlich ein Wortartenwechsel vorgenommen werden hin zum Adverb „englisch“ vorgenommen werden müsste. Im Kontext der Gesamtaussage des Kommunikats ist das Deutsche fehlerhaft.

6 INTRODUCTION

This volume of the manual aims at compensating for the workshop that was scheduled for spring and that had to be cancelled due to the Corona crisis. Its focus is on multimodality again. The second part of the manual shared in spring was dedicated to the fundamentals and offered a first insight into the functioning of multimodal artefacts (multimodal aggregates). This third volume of the manual builds on the findings of volume two and turns to multimodal practise. Examples of different complexity and thematic key aspects show how language and image interact. The examples not only address the relationship of text and image but also different forms of background knowledge needed for understanding. This knowledge is either expected to exist in the perceiver or is furnished by the multimodal artefact itself. The contents of this manual wants to shed light on the composition of multimodal aggregates, to support their understanding and furnish help for the interpretation of data.

Chapter 7 is dedicated to the acquisition of multimodal aggregates and pertinent the cognitive processes. Chapter 8 connects the results from chapter 7 with visual training and the interaction of cognitive processes and conventionalization. Chapter 9 continues this perspective and examines the performance of multimodal data. In this chapter those examples that do not stem from advertising give proof of the scope and variance of multimodal aggregates. The final chapter 10 then turns to the limitations of different modalities and discusses the impact this has on the construction of multimodal aggregate.

7 TEXT AND IMAGE

The analysis of multimodal aggregates is part of image linguistics, which is an independent linguistic sub discipline that can be located at the interface of sociolinguistics, text linguistics and media and communication studies. Linguistic landscape research only marginally considers the topics relevant for image linguistics.

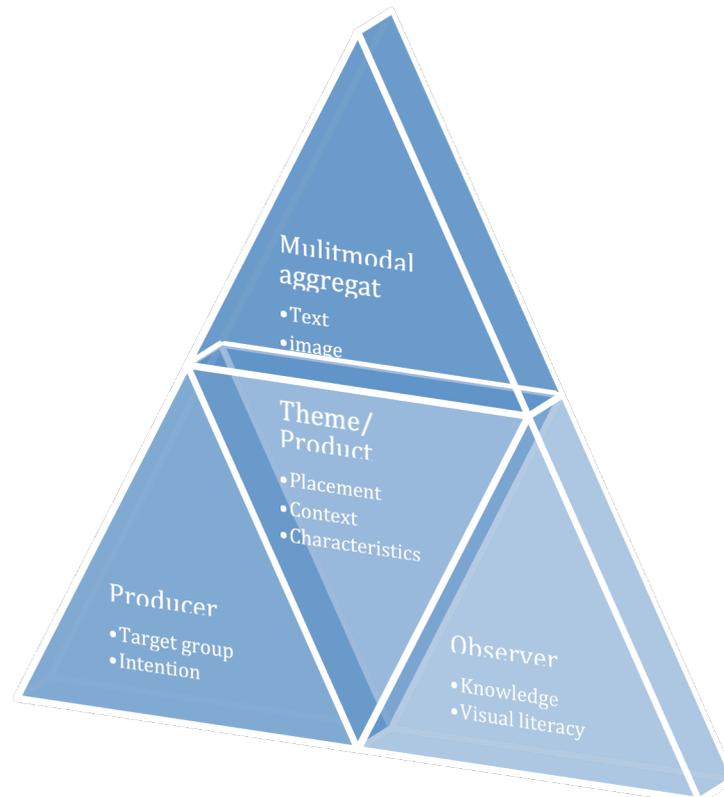
An important discussion from image linguistics is that about the definition of the term "text". In a standard sense this term only refers to a purely linguistic thing that has to fulfil certain requirements with regard to form and content. In image linguistics and multimodality research, however, there is the need to examine the contributions of additional modalities to coherence and cohesion in propositions exceeding a single sentence and having narrative or descriptive character and power.

At the level of non-linguistic signs code has to be distinguished from mode, though sometimes they are used interchangeably. Colour is a modality, yet there are colour codes. The latter refer to conventionalised meanings of certain colours: black as colour of mourning, red as colour of love, white as colour of innocence etc. These conventions endow colours with symbolic meaning also known for flowers, e.g. red roses as symbol for love, gladiolus as symbol for pride.

The following questions can/have to be considered either as independent research questions or as part of the analysis of multimodal aggregates:

- Which kinds of sign contribute to the construction of a multimodal aggregate?
- How does a sign contribute to the construction of a multimodal aggregate?
- Which is the part of
 - Producer
 - Which target group does he/she aim at?
 - What is his/her intention?
 - Perceiver/Observer
 - Which are the perceiver's capabilities to decode a multimodal aggregate?
 - What may arouse the perceivers attention
 - Product
 - The situation in which the observer perceives the product
 - Explicitness of the aggregate

A central aspect is world knowledge. The sign producer relies on the perceiver having a certain body of knowledge comprising a society's conventions and a certain degree of factual knowledge etc.. The sign more or less directly refers to this knowledge which, upon perception, is activated in the perceiver and which hopefully results in understanding an aggregate the way intended by its producer.



7.1 Acquisition of multimodal aggregats:

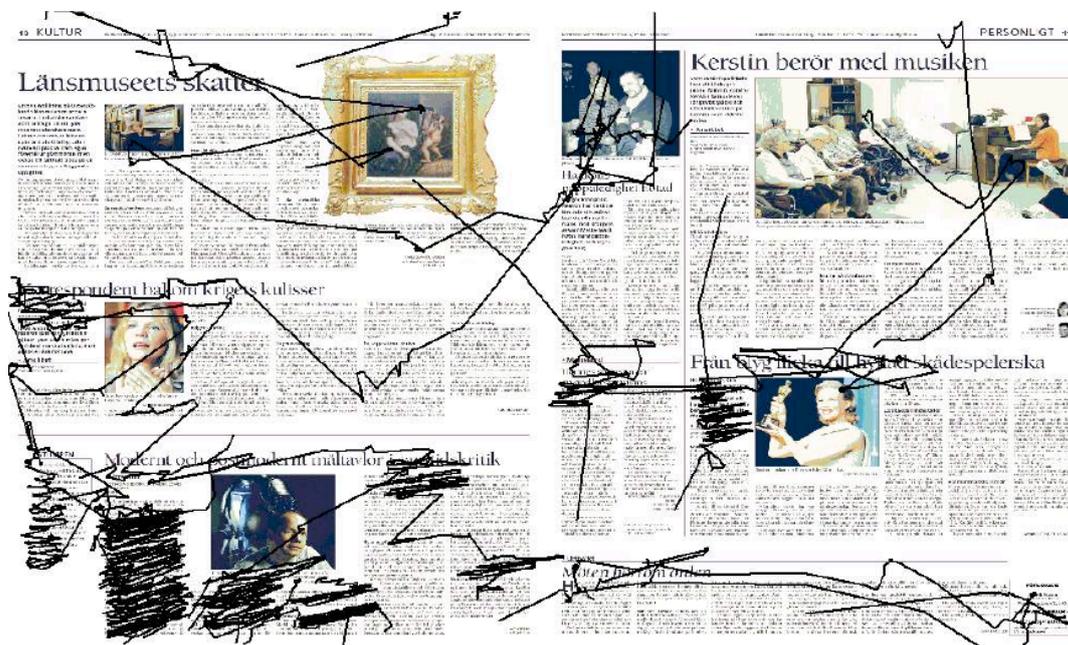
Depending on the context in which an aggregate is perceived, topic and/or local contexts as well as personal experiences and interests can influence the de

The construction and acquisition of multimodal aggregates is influenced by two major general cognitive rules:

1. Images are always looked at
2. Images showing faces are looked at more intensively and longer than images

These cognitive imperatives show that images have an evolutionary advantage over text as the product of writing which is a cultural skill but no cognitive or other human universal. In sum the cognitive factors result in various possibilities for a perceiver to construct pertinent units in a multimodal aggregate. The examination of how perceivers form units is carried out using eye-tracking methods.

The first example, taken from Holmqvist & Wartenberg (2005), shows how a newspaper broadsheet is perceived and which items are looked at. Eye-tracking results show that all images are viewed at least shortly. The same holds for the headlines, which are looked at more intensively than other pieces of text. This can be explained with their sizes, which draw attention to the headlines and make them more salient than the running text. However, this judgement that influences a perceiver's behaviour is learnt knowledge regarding the function of salient pieces of texts in newspapers.



Holmqvist & Wartenberg (2005: 6)

Figure 4: Reading pattern over a broadsheet spread recorded as absolute data.

The second example is a so-called complex picture with motifs from a Swedish children's book. This kind of complex picture simultaneously shows activities, which are normally carried out successively. In the picture shown here it is preparing a field for sowing and the act of sowing (Boerris & Holsanova 2012; all images below from Boerris & Holsanova 2012 in Bateman 2014:247).

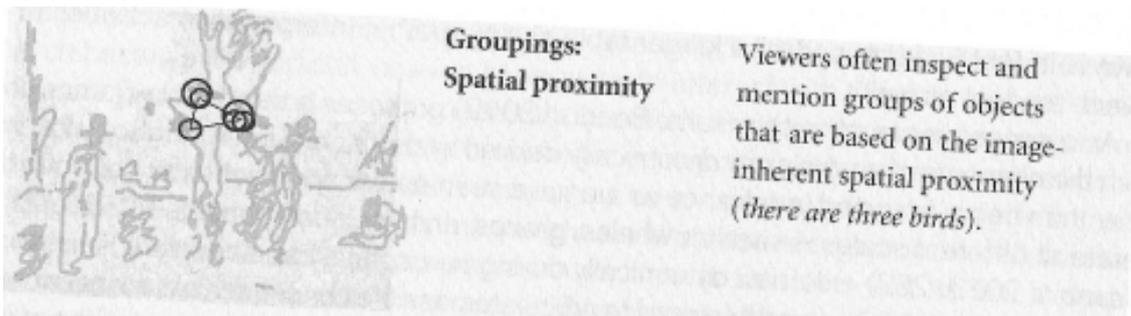
Pictures of this kind are used to examine the formation of aggregates, i.e. the way in which objects in the picture are related to other objects present. Research shows that there are two mechanisms involved, which are both based on the cognitive process of categorization and that result in prototypical ways of grouping.



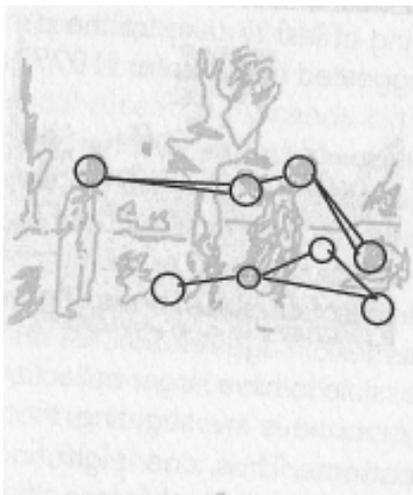
Figure 3 Complex picture: the motif comes from a children's book by Sven Nordqvist (1990). Reproduced by kind permission from Sven Nordqvist.

The order and intensity of perception are influenced by

Spatial proximity



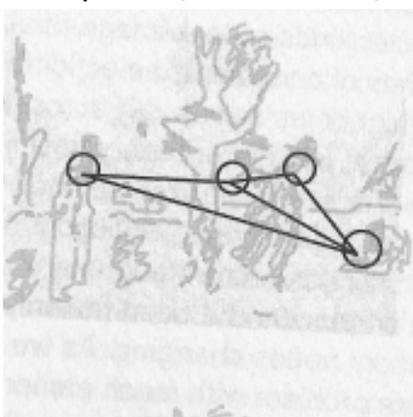
Categorical similarity (i.e. all persons, all animals, all trees etc.)



**Groupings:
Categorical
similarity**

However, viewers also perceive and describe clusters of objects that are *not* spatially close. These clusters consist of groupings of multiple similar objects distributed over the whole scene (*four cats, four versions of Pettson*). This type of cluster is based on categorical similarity.

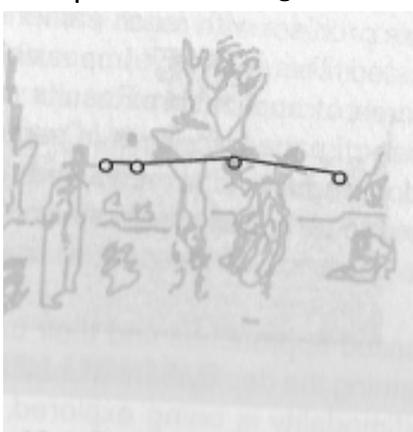
Multiple objects (i.e. objects that are present more than once)



**Multiple objects:
Simultanbild**

Some viewers notice repetitive figures in the scene and describe them as four varieties of Pettson and Findus, whereas others focus on the event instead and describe one person and one cat involved in various activities.

Composition along lines (e.g. horizon)

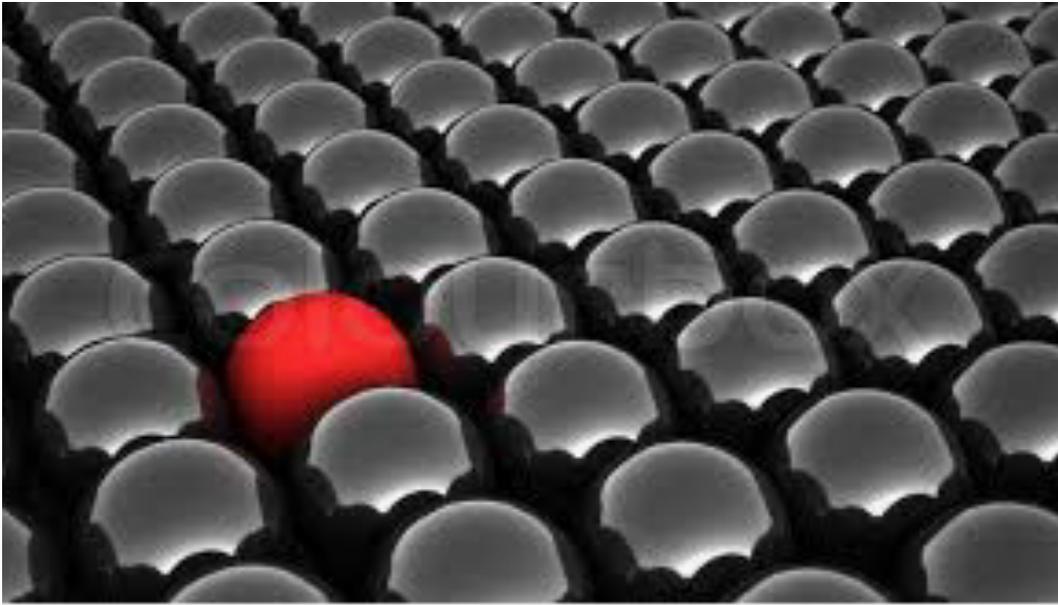


**Grouping:
composition**

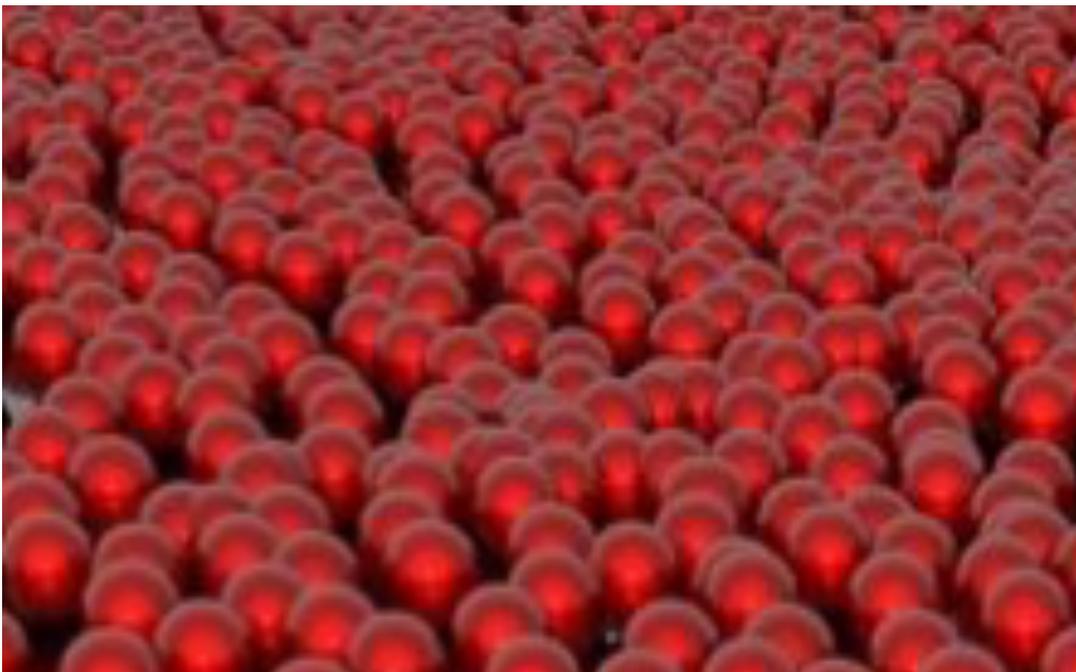
Another type of cluster perceived as a meaningful unit in the scene is *hills on the horizon*. The viewer's eyes follow the horizontal line, filling in links between objects. This cluster seems to be a compositionally guided grouping.

This form of categorization has evolutionary and cognitive foundations that influences visual attention. Looking at interesting objects first and fast has evolutionary reasons such as quickly detecting prey, enemies or dangers. This results in stimuli being especially eye-catching (salient) in certain contexts. In the four following pictures the

red ball and the crocus are only salient in the context created in the pictures. The red ball does not arouse attention if places among other red balls and neither would the crocus be salient among other crocuses of the same colour.



Boeris & Holsanova 2012



<http://bilder.4ever.eu/zeichentrick/digital-art/rote-kugeln-183874>



Boeris & Holsanova 2012



canstockphoto.at

7.2 *Distinguishing ways of perception*

Basically there are two different forms of visual perception:

1. **Bottom-up:** This process is salience-based. It is peculiarities that activate and direct perception (see above examples)
2. **Top-down:** This process is conscious and task/interest specific. The perceiver actively looks for a certain piece of information. The perception of objects and their relevance to the perceiver depends on the latter's task or interest.

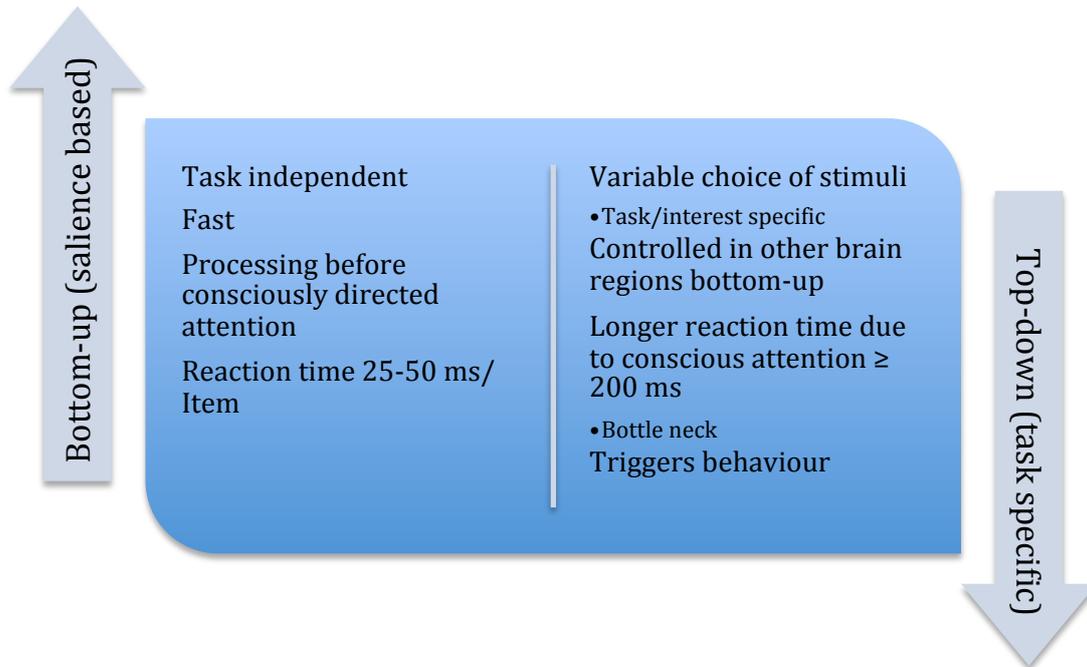
These two forms of perception can briefly be characterized as follows:

Salience based is

- Independent of the kind or content of a given task
- Fast
- Mainly bottom-up driven
- Probably processed before consciously directed attention is activated
- Reaction time is 25-50 ms per salient item

Top-down as form of attention is more conscious and powerful

- Variable choices for selecting stimuli
 - Task and interest specific
- Probably controlled in other brain areas than bottom-up processes
- Conscious attention results in longer reaction times ≥ 200 ms per item
- Causes bottle neck
 - No complete simultaneous processing of data
 - Serial processing
 - Disintegration of complex visual scenes into a fast series of simple and narrow analysis
 - Triggers behaviour
 - Closely connected to recognition, planning and movement control



Both “techniques” are used with all visually perceptible objects:

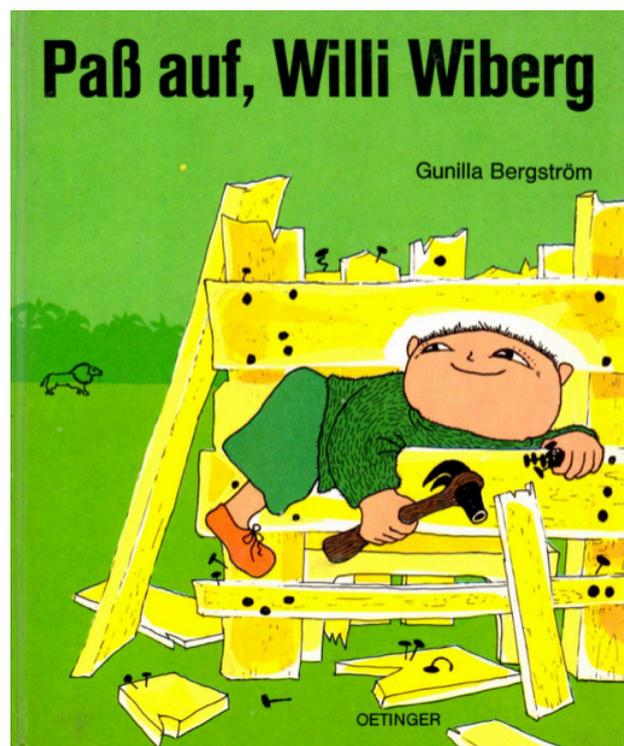
- Image only
- Text only (also including font, size, bold, underlined, italics, colour etc.)
- Combinations of text and image

7.3 Text-Image combinations

Knowledge of cognitive mechanisms that control visual perception supports the conscious composition of multimodal aggregates. In these aggregates images have special functions with regard to the aggregate’s overall meaning. Correctly interpreting images in a multimodal context is based on multimodal literacy. Multimodal literacy is learnt from childhood on and is considered a core competence when using modern media. To demonstrate the degree to which visual perception is trained and shaped from childhood on, the different functions of images will be explained with the help of the cover of a volume of the Swedish children’s book series Willi Wiberg.

Images do a lot more than just illustrate text. They

- Construct a scene
 - The cover shows Willi doing something that is typical for the story told in the book
- Define/develop characters
 - Willi obviously sometimes does things, he should not do. Willi is a preschool boy and is using a hammer and nails. The smile on his face shows that he likes doing forbidden things.
- Expand/develop the plot
 - Willi is constructing something. He is involved in an activity that is potentially dangerous for children but that is fun.
- Change perspectives
 - The title represents the perspective of parents. The smile on Willi's face stands for another opinion.
- Contribute to the coherence of the text
 - Text and image are closely connected.
- Emphasis the text message
 - Hammer and nail can potentially be dangerous. One should be careful when working with them. This fits the title.



Not all graphic/pictorial elements directly contribute to the information infrastructure:
The following functions can be distinguished

- Decorative function
 - Invitation to a children's party → balloons
 - Little to no direct relationship with the text
- Representative images
 - Accurately reproduce the entire text or parts of it
 - Most frequent form
- Organisational images
 - Structuring of the information given in the text
 - Steps of a procedure
 - Maps
- Interpretative images
 - Visualisation of complex systems or facts
 - Blood circuit
 - Technical drawings
- Transformative images
 - Frequently mnemonic function
 - Graphic representation of abstract facts and relationships
 - Graphic cribs
 - E.g. representation of the place name "donkey castle" with the help of a picture of a donkey and a castle:



© Bilder: Wikipedia

8 VISUAL TRAINING

Visual training is the acquisition of ways of perceiving and interpreting non- and paralinguistic information that is visually conveyed. Visual training is based on prevailing societal conventions and can also reflect fashions or historic values. A change of values can be a source of misinterpretation or conflict. The same holds for encounters of different societal/cultural groups, see e.g.

- Cultural function of the colour white
 - Europe: Joy, purity
 - Parts of Asia: Colour of mourning.

These conventions are so strongly internalized that it is necessary for internationally or globally active companies to localize their advertisements etc. Localisation is the culture specific adaptation of advertising contents in text and image. As globalisation has brought about a more general knowledge base with a strong western influence, it is not necessary to develop individual advertising campaigns for each target region but only to adapt them in major critical points.

A simple example is charging fonts with certain associations.

- Fonts with serifs convey motions of authenticity and tradition
- Fonts without serifs convey reason and classiness

These associations influence the choice of fonts for brand names and advertisements:

The following two labels want to express different things with their choice of font. While Captain Morgan uses a font with serifs in order to express an idealized imagination of the independent characters and way of living of pirates, the Whiskey label wants to emphasise tradition. All these associations can be attributed to both the image of the label and the customers. Thus the associations the fonts activate influence the way the label and the customers are perceived by others. This interrelation between label and customer is always part of the image of the label, which not only is to convey a certain image of the brand but also aims at the customers identifying with that image and transferring its notions and association onto themselves.



The following examples show the use of fonts in up-scale labels. The famous Chanel logo uses a simple and clear font without serifs. This reflects a certain perspective on luxury, i.e. the “equation” that classy/elegant is associated with an appearance that is lean, clear and simple at first sight and the sophistication of which is only visible for insiders.



The label Bulgari chooses a slightly different approach to its image. It uses a font with modest serifs. However, it does not want to activate the above mentioned associations. It is the chosen font that makes distinction possible. The font corresponds to the classical Latin alphabet that is associated with ancient Roman inscriptions. The essential feature is the use of “v” instead of “u” that demands a certain knowledge in the perceiver in order to understand the sign. Additionally, the

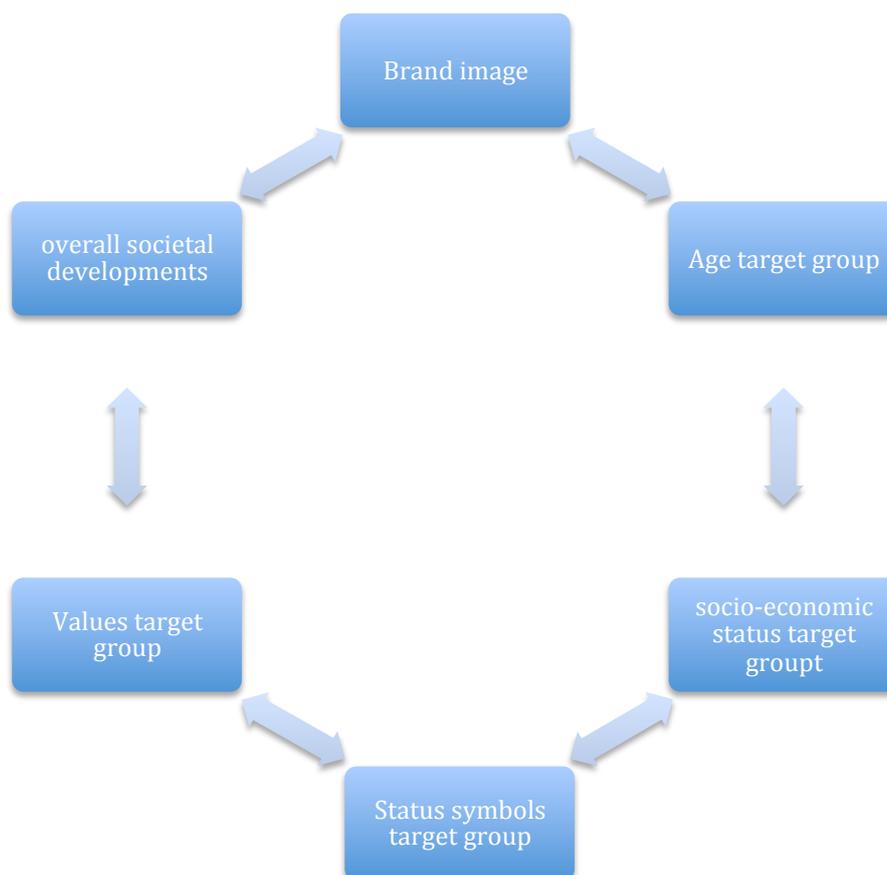
reference to Rome can activate to associations: The first is that of the Italian capital, where the company is headquartered. The second association is that of the antique place that is known for the luxurious way of living of its upper classes. This association is possible only because there is little knowledge about the everyday life of the ordinary Roman citizen in present day public.

The French luxury label Hermès has a similar approach but uses a different font. The company name is presented in a font with serifs. The notions of tradition and authenticity associated with this font makes reference to the roots of the company in saddlery. This connection is also present in the graphic that is part of the logo. This depiction however, does not only address the original type of business the company was involved in, but also the wealthy class of customers. This reference to wealthy customers (it only those who could afford the carriage and clothes depicted) is mirrored in the font without serifs chosen for the placename. Paris is not only associated with wealth but also with elegance and fashion, which is indicated with the choice of the font. In sum there is a complex logo with the fonts representing certain values and the depiction, though placed in the upper part, unites the messages of both fonts.



It is important to keep in mind that fonts, colours etc. are always processed on the basis of learnt associations. This happens unconsciously and automatically as those associations are part of an supra-individual, overall knowledge. However, as there is awareness as to the existence of this knowledge, producers consciously choose colours and font to align with societal convention when constructing a brand image. This is the only way to make sure that identification with the brand is possible for the target group. Introducing a new brand thus involves determining a target group and its values, which have to be considered when constructing the brand image. After introduction into the market this target group must be observed closely in order to adapt to changes in its values etc.

In case e.g. a brand for young people wants “to grow up” with its target group, advertising and brand image have to be altered accordingly. In case of status symbols of a target group becoming obsolete the entire brand image has to be adjusted, as it heavily relies on the wilful activation of certain associations which trigger categorization (see above), which places the brand and its customers in certain economic and/or societal segment:



The following example of the Café of the company Riquet from Leipzig shows how the presentation in public does not only represent the company but also makes reference to fashions of the time of the construction of the company image that were important when addressing the target group. On the company's website the places architecture that appears out-dated, kitschy or even colonial is connected to its date of origin and thus transformed into a historical statement.

Herzlich Willkommen im Riquet-Haus

Das 1908/09 im Auftrag der Firma Riquet & Co vom Architekten Paul Lange errichtete Geschäftshaus gehörte seinerzeit mit zu den modernsten und zweifellos originellsten Neubauten Leipzigs.

Die bis ins Jahr 1745 zurückreichende Handelstradition der Firma Riquet mit Ostasien und dem Orient setzte der Architekt auf einfallsreiche Weise um.

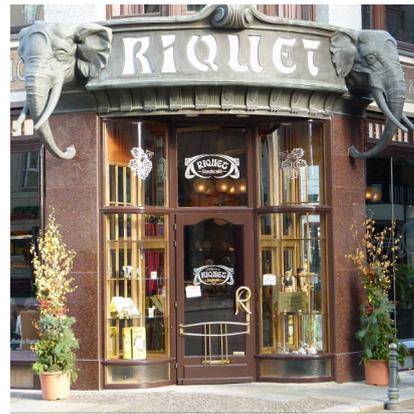
Das doppelstöckig geschweifte Dachtürmchen entlehnte er der klassischen chinesischen Baukunst. Brüstungen, Pfeiler, Hauptgesims und Rundgiebel sind mit wunderschönen farbigen Mosaiken in Jugendstilmanier geschmückt. Zwei, die Eingangstür zum Cafehaus flankierende, kupfergetriebene Elefantenköpfe stellen das Markenzeichen der Firma Riquet dar.

Welcome to the Riquet Building

The house was built in 1908/1909 by the architect Paul Lange on behalf of the company Ricquet & Co. After its completion it was one of the modest modern and fancy new buildings of Leipzig.

The architect planned the building to give proof of Riquet's trading history with East Asia and the Orient that dates back to 1745.

The two-floor curvy roof tower was taken from traditional Chinese architecture. Railings, pillars, fascia and round gables are decorated with colourful art nouveau mosaic. Two embossed copper elephant heads flanking the entrance are the hallmark of Riquet.



<http://www.riquethaus.de/kaffeehaus.php>

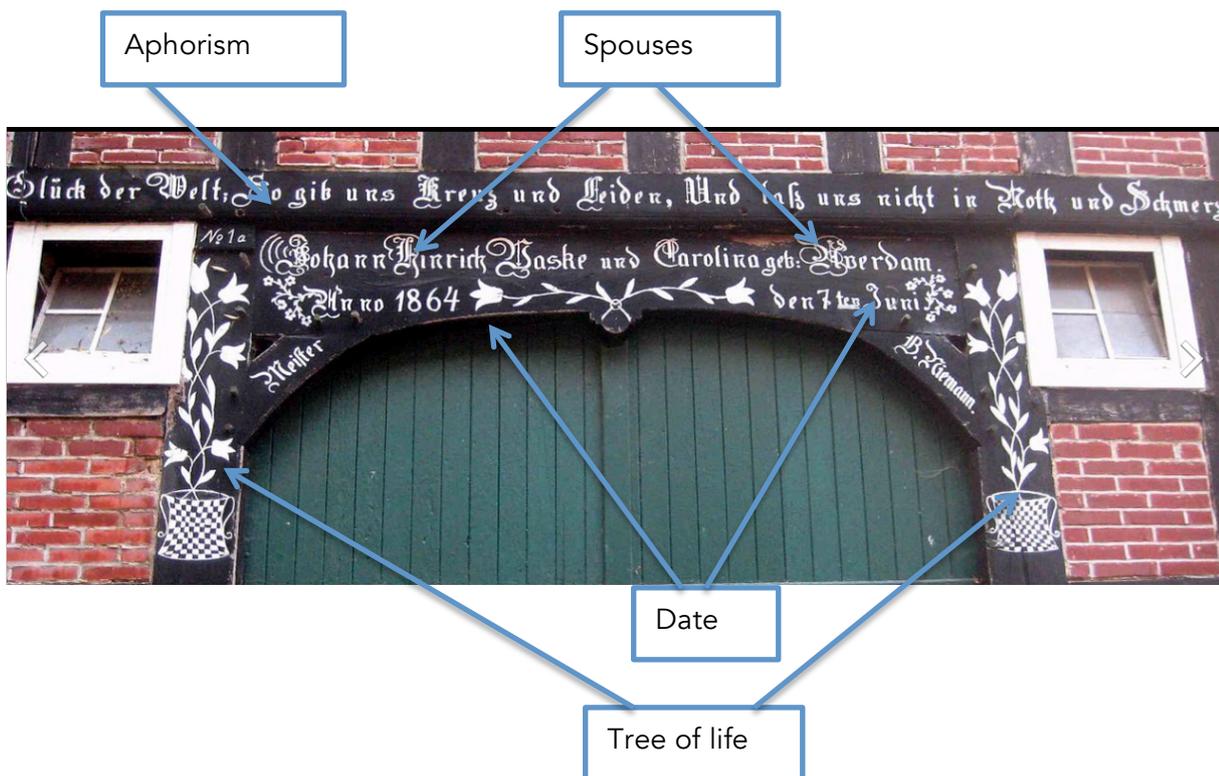
This example also shows why it is important to consider architecture in multimodal analysis of public space. Integrating architecture can make reference to different aspects: It can reflect a company's history (as is the case with Riquet) but also mirror more general societal and/or political development and events at different levels (history of families, places, local myths, traditional handicraft etc).

9 IMAGE, COLOUR, LANGUAGE – THE INFORMATIONAL LOAD OF MULTIMODAL DATA

Multimodality always is the simultaneous presence of pieces of information that are presented in different semiotic systems. The combination of these single items forms the overall information, which are more than just the added meanings of the individual pieces of information. This approach helps to extend the scope of analysis of classical linguistic or semiotic landscape data und offers possibilities the examine the use of languages and different discourse types in public space. This chapter shows occurrences of multimodal practise in public space. Different examples show how multimodality works and which aspects have to be considered in analysis.

9.1 Multimodal aggregates as mirror of social structures and traditions

The sample in this section shows that cultural practice can be a part of public space. In some regions in Germany it was a local tradition to decorate the house of the newly wed in the way shown in the picture. The house is decorated with a religious aphorism, names of the spouses, the date of the wedding and two trees of life that symbolically represent the foundation of a new family, which is to prosper and flourish.



A perceiver not knowing the underlying traditions will neither be able to dissolve the aggregate nor understand its meaning. An ignorant perceiver might assume that there is no connection between the date given and the other items. The trees of life can be interpreted as mere decorative floral paintings. The year given might be year the house was built and the names could be the names of the building owner.

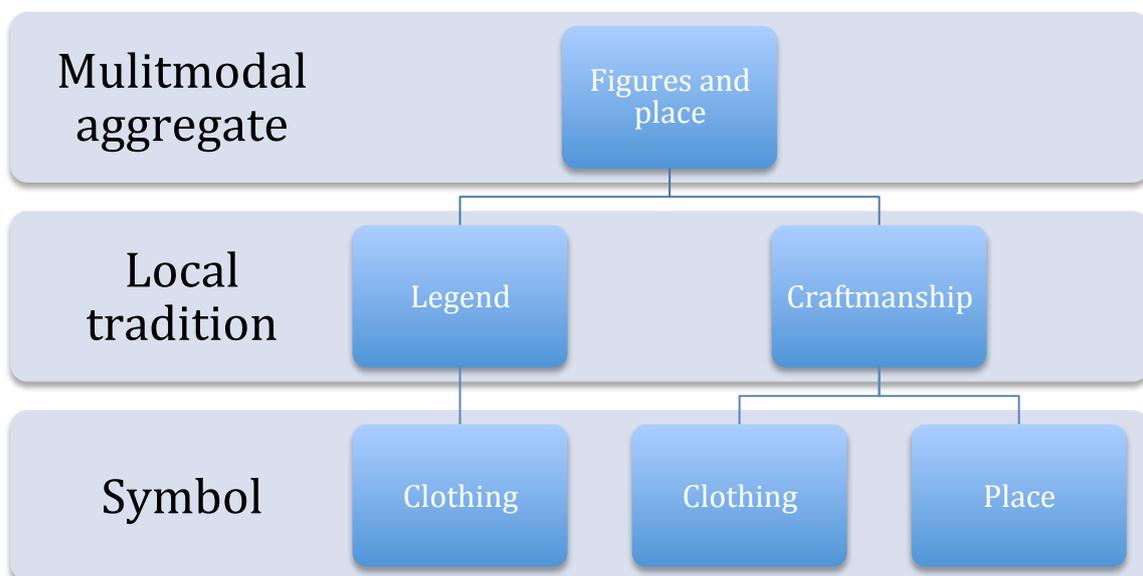
Yet, this example combines language and culturally charged symbols, including the house in the sense of home to one statement, turning the private public thus integrating it into the valid social norms and expectations. In the linguistic landscape of a place these forms of public display did not only represent historic values but also structured the place, publicly established relationships to other inhabitants (relatives) and can, if existing in an adequate number, publicly display the place's social network and contacts.

9.2 Language free presentation of local myths and traditions

The second example shows the integration of local myths, traditions and history into public space. The picture shows the intersection of Schäfflerstraße and Weinstraße in Munich. This place is called "Schäfflereck" (cooper's corner) in common speech.



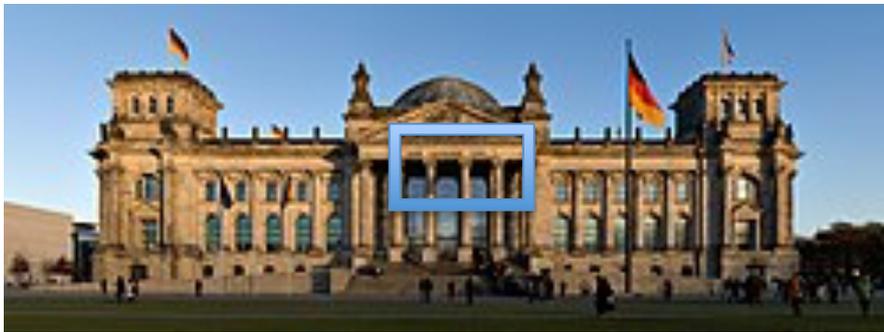
The position of the figures is a reminiscence of historical structures. Street names that consist of names of handicrafts normally indicate the quarters where these crafts traditionally were located within a town or village. Hence the figures at the beginning of Schäfflerstraße indicate that here the coopers of Munich used to have their homes and workshops. What is special about the figures is their clothing, which is not their traditional working clothes, but their dancing clothes. Showing the coopers in these clothes takes up a local legend according to which the coopers of Munich publicly displayed dances in the city in 1517 during an outbreak of the black death. The coopers wanted to morally support the people, entertain them and make them leave their houses and take up everyday life again. Since 1760 the coopers' dance has been performed every seven years. Further references to the coopers' dance can be found in the city centre of Munich. It is irrelevant that the historical correctness of its origins of the coopers' dance is questionable as it seen as a legend.



The figures do not display any language. They draw their meaning and justification from their location that is closely connected to historic structures of Munich thus identifying the figures. The figures themselves and especially their clothing hint at an event in Munich history that is not connected to their handicraft. This way the figures tell a story about a historic event in Munich that turned into a local legend without using any language.

9.3 Text-Architecture Aggregates

The third example has another scope and is part of the complex of historic monuments. Despite the examples one and three also having a historic dimension, they cannot be compared with this example as they are local and social in meaning only and lack the "national", political dimension.

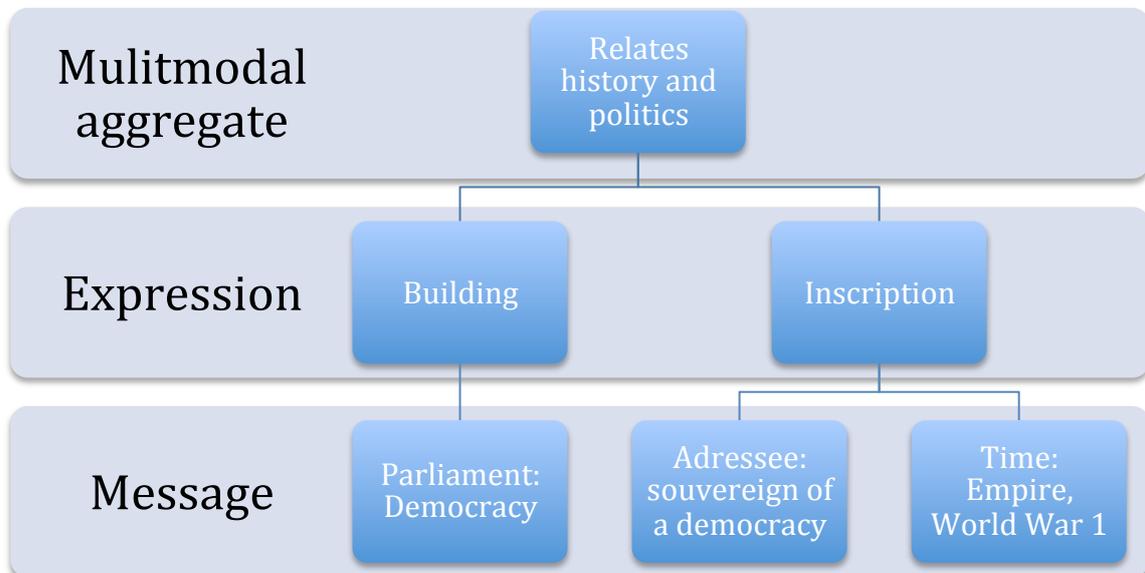


Bilder: https://de.wikipedia.org/wiki/Dem_deutschen_Volke

The inscription is located above the western entrance of the Reichstagsgebäude in Berlin. The inscription dedicates a central institution of democracy, i.e. its parliament to its sovereign, the people. The inscription was placed above the parliament entrance in 1916. On the one hand the inscription's positioning provides an air of timelessness that at the same time is challenged by the reference to World War I and the German empire activated through the inscription's year of establishment.

The inscription's central meaning is produced by the interplay of the building and its function. Placed anywhere else the inscription would appear odd or even not understandable. The meaning is supported by the flags on the building, which indicate its special public and national character that is thus discernable for any observer without any special knowledge, as displaying flags on non-official buildings is not common in Germany.

This example combines writing, location and additional symbols, with each element explaining the presence of the other elements. So from inscription and flag one can conclude the character of the building, which again is a precondition for the presence of the text and the flag.



9.4 Function and meaning of non and paraverbal elements

This example is taken from a comic and shows that the literacy of multimodal aggregates is learnt from childhood on. This refers to the interpretation of non-linguistic elements as well as to the meaning construction from the interplay of non-linguistic and linguistic elements.

The importance of non-linguistic elements becomes obvious when imaging the images without them.

Bold letters, sound waves

Position persons

Different kinds of lines indicating movement

Onomatopoeica

Language free symbol and known action

Different kinds of lines indicating movement

Panel 1 (top-left): **PACHT IHN, IHR FEIGLINGE! ODER ICH WERF' EUCH IM ZIRKUS DEN LÖWEN ZUM FRASS VOR!**

Panel 2 (top-middle): **Den Löwen?**
Zum Fraß vor?
Als-dann...

Panel 3 (top-right): **Was ist denn los?**
Ein Gallier hat sich ins Lager eingeschlichen...

Panel 4 (bottom-left): **KLICK!**
TSCHACK!
KLONK!
KLICK!

<http://www.erzaehlenbilder.de/2015/01/06/zack-bumm-knuff-ganz-einfach-oder-doch-nicht/>

Bold capital letters = shouting

Semiotic support through the picture.

- Sound waves for volume
- Clothing: speaker = Officer
- Scared = is hiding behind a tent



The different modalities interact in the following ways.

1. Interaction of text and typography

The text expresses an order. This becomes obvious through wording and the way it is uttered. It is no polite request, but it is shouted. Shouting is indicated with bold capital letters. This representation supports the wording. Typographic aspects also contribute to the correct interpretation of the statement and emphasise or support it. The aim is a semiotically supported imitation of the speech act.

2. Nonverbal characterization of the speaker

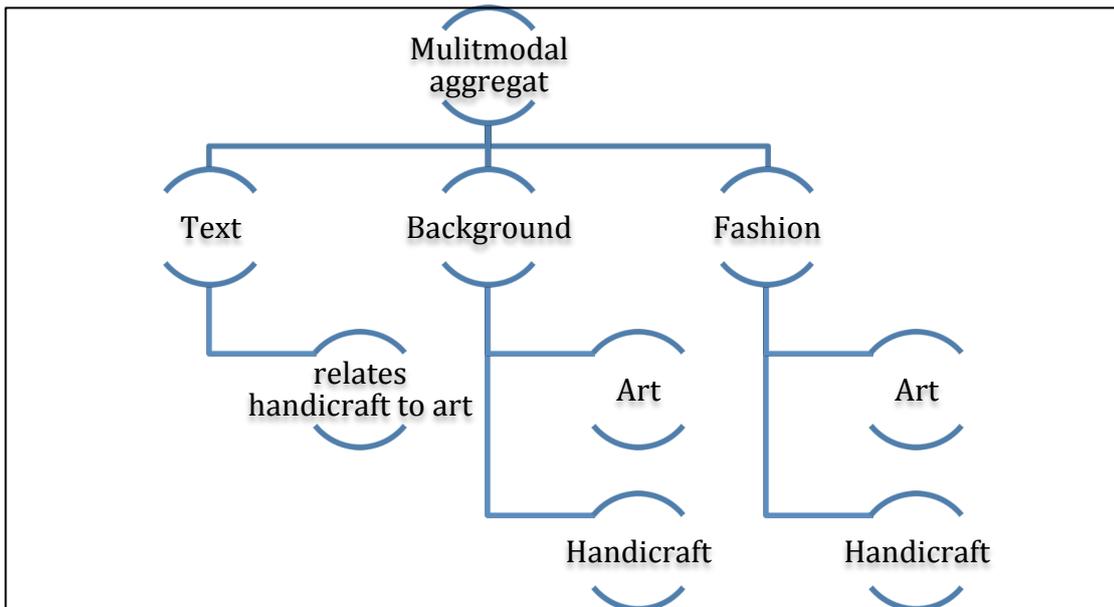
Clothing identifies the speaker as an officer who has the authority to give orders. The position of the figure constructs a slight contradiction as it indicates that the officer is scared of something, hides from this danger but exposes his subordinates to it.

9.5 *The interaction of components in complex multimodal aggregates*

This example shows how complex messages can be constructed in public space using the multiple levels of shop windows.

The text on the shop window deals with art, handicraft and nature as well as the artful refinement of the latter to turn it into art. The text refers to the background of the shop window, which is not only some sort of decoration but art on the basis of lotus leaves. The depiction of a lotus leaf on the upper right corner of the text establishes the relation between the text and the piece of art. The clothing displayed is placed in the middle of the shop window and thus between text and piece of art so as to transfer the notion of the text and the background not to clothing but to tailoring. In this context it is important to know that in German a word for handicraft is "Handwerkskunst", i.e. it is a compound of "handicraft" and "art". The word "art" can be added to most names for crafts in German. So tailoring can be "the art of tailoring". The text is legible only at short distance. From a larger distance the complex interpretation of the show window is not possible; all that can be seen then is the clothing against a green background.



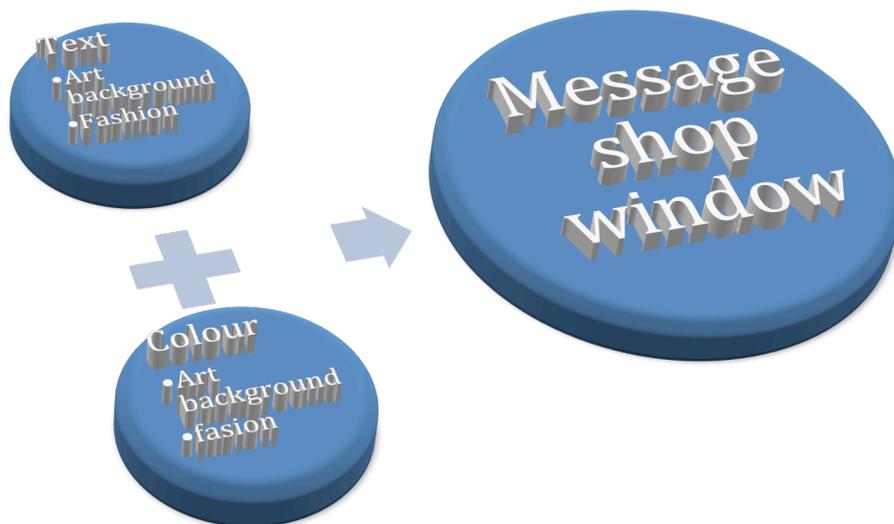


The same composition is used in the neighbouring shop window. The only differences are the colour of the piece of art in the background and the type of clothing displayed. So, here too, proximity to the shop window is necessary to detect and decode the complex messages. From a distance it is again only a background with no specific meaning and the clothing that dominate perception.



Both decorations use a contrast of background and clothing displayed to ensure good visibility of the latter. Another point that requires attention is the relationship between background colour and the style of the clothing displayed. In the first window there is a green background. This colour is prototypically associated with nature, which fits the kind of fashion displayed (young people and leisure). The dummies are grouped around something that resembles a bus stop and that might indicate activity support this notion. In contrast, the second window shows formal clothing against a golden background constructing a more reputable atmosphere.

Both windows show that addressing art and constructing a connection to craftsmanship upgrades the products displayed as the notion of art is transferred to them. Those observers of either shop window who do not perceive the writing still can profit from the interaction of colour and style of clothing, which a further elements supports in the first window.



9.6 *Nonverbal emphasis of a topic*

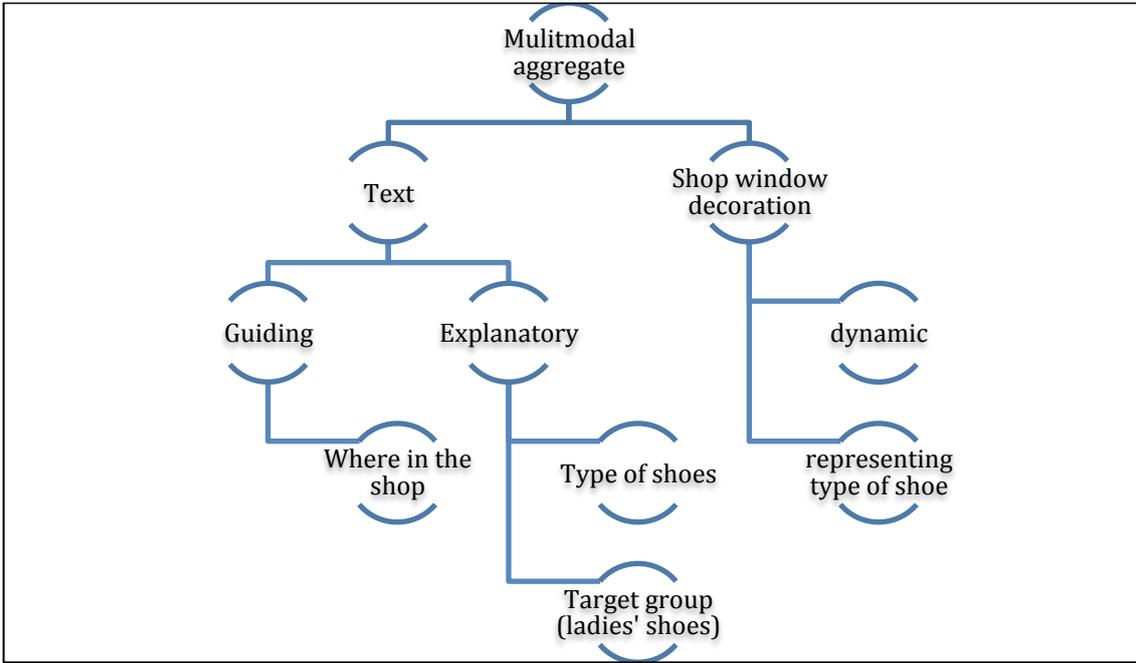
This last example shows how the topic of a multimodal aggregate is taken up and extended nonverbally.



The linguistic part in the shop window is explanatory and guiding

1. The explanatory character is furnished by "sneakers women" as the products displayed cannot unequivocally be identified as sneakers for women but could also be models for men and women alike.
2. The hint "1OG" (1. floor) leads the way and informs the observer where to find the shoes inside the shop.

The arrangement of the products is symbolic with regard to the classical area of application of sneakers. The sloped shelves have a dynamic character and indicate movement. Even if sneakers today are part of certain fashion styles and are not only worn during sports activities, they are still associated correspondingly.



10 IMAGE, COLOUR, LANGUAGE – THE LIMITATIONS OF SEMIOTIC SYSTEMS

The previous chapter focused on the informative value of multimodal aggregates and the interaction of modalities. The construction of multimodal aggregates is not only subject to learnt conventions but also to the limitations of each semiotic system. Not every modality can express anything. The following examples offer a first insight into the limitations of semiotic systems and thus strengthen the understanding for the composition of multimodal aggregates.

The limitations of the single modalities derive from two different components.

1. Conventionalized knowledge. In the following example this is e.g. the meaning of colour.
2. Cognitive processes
 - a. Images are perceived holistically i.e. in their entirety at one glance. The part of an image that serves as “entry” to a closer inspection depends on context and is either salience based (bottom-up) or task/interest specific (top-down). Due to their special way of processing, images are memorized faster and better than language
 - b. The perception and processing of language/text is always sequentially (word per word). Thus, the order cannot be chosen deliberately as it will hinder/prevent understanding.

10.1 The limitations of colour

The previous manual described colour as a dependent modality as it does not have meaning on its own. The meaning attributed to a colour is object and context specific. Additionally colour has a conventionalized, culture specific and thus varying meaning as well as a more general meaning that is derived from its occurrences in nature.

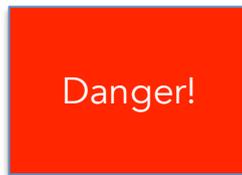
Reference to love is a socially conventionalized meaning of the colour red whereas attracting attention and warning are functions that have their origins in the signal system of nature. Thus the red square below does not have a specific meaning and can be used for all three purposes mentioned. So, it is the text that provides the bases for the final case specific meaning of red.



More than one possible meaning



1. Attention



2. Warning



3. Emotion

The form of interaction as shown in the examples can also be transferred to the interaction of a text and the colour of the font it is written with. This interaction is illustrated with an example already presented in the last volume of the manual. The red colour used for writing does not only add meaning to the word but also adds the notion of love to the items displayed in the shop window.



10.2 The variance of meaning of images

As is the case with colours and other dependent modalities the meaning and message of images is not unequivocal so that the same image/picture can be used to convey different meanings. The below picture of a wildflower meadow will serve as an example:



In this first example the wildflower meadow only is the background for the main object, which is the bee asking for rescue. This use could be part of a campaign for environmental protection. The meadow is then part of nature that deserves protection and that depends on the existence of bees. However this implication is not addressed directly so that the meadow primarily is conceived of as the living space of the bees.



The second picture is used in a similar way and is only indirectly connected to the topic, i.e. owning a house in the country. Yet „house in the country“ also refers to country life and idealized ideas about country houses with gardens and untouched nature. Thus the picture serves as representation of this idealized ideas about living in the country and is the end of a chain of associations, which are triggered by the text.



In the last example the meadow is given the contrary meaning as the text produces a context in which the meadow appears as the source of disorders.

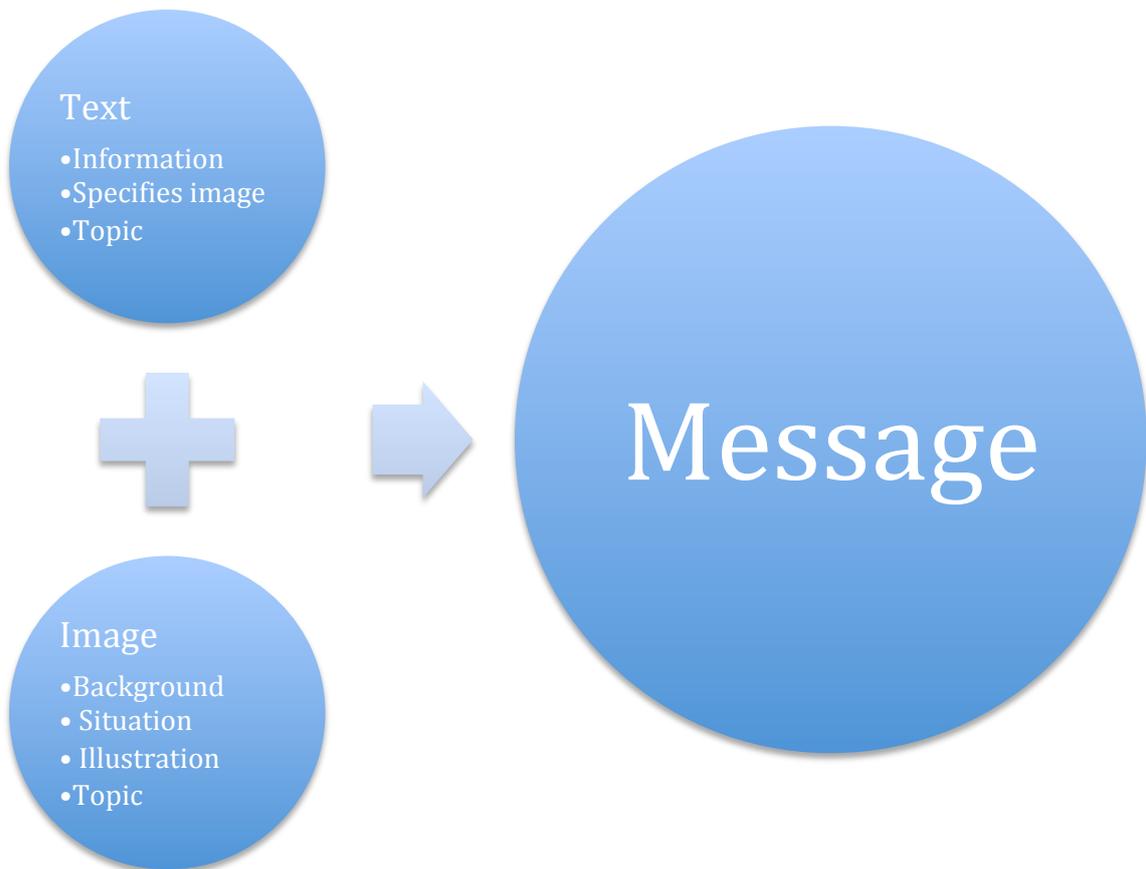


The feature the two last cases share is that a relationship can be established between the picture and at least one word from the text. While in example 2 there is a more or less obvious relationship with the word "country", the connection with the word "summer" might seem a bit stronger.

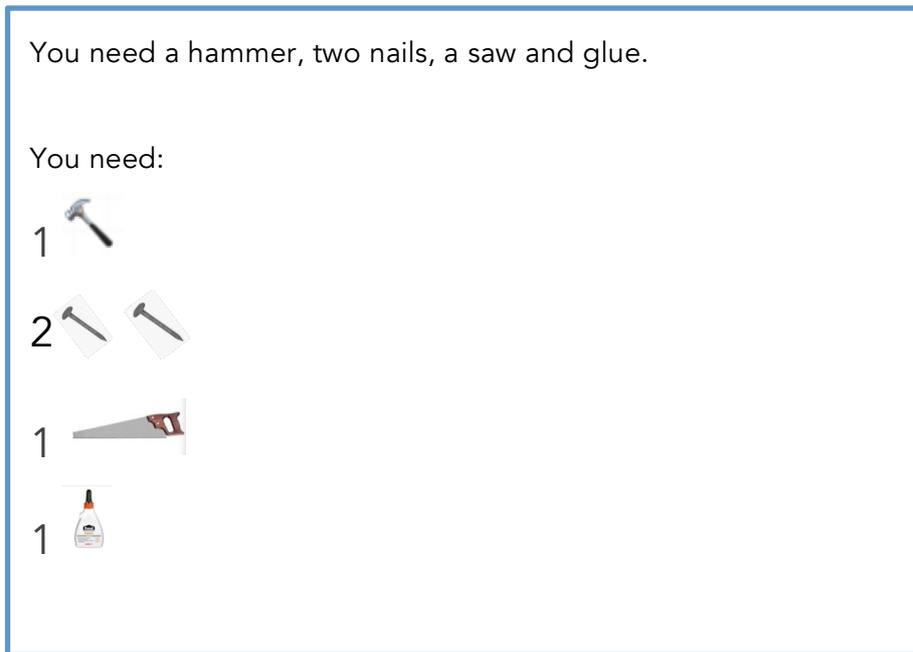
As is the case with the examples for the colour red, here too, it is obvious that the meaning of images, graphics depends on the context and can, depending on the construction, have varying influence on the overall message of a multimodal aggregate. All three examples can also be imagined to be constructed with an image other than a flower meadow. Thus, choosing an image/graphic in advertising not only depends from the message that is to be conveyed but also from the image of the brand and the target group addressed.

The entire message can also be conveyed with text only. In this case processing time would increase as a text describing all notions the picture conveys would be rather long and would have to be read and processed sequentially. At the same time such a text would be less emotional as an image as the latter immediately activates associations, which are much more direct and which additionally connect to the personal experiences and knowledge of the perceiver, compare Straßner (2002) who describes pictures as "quick shots into the brain".

Below picture is a last example on the emotionalizing and contextualizing function of images. The picture in the background and the Vespa in the foreground emotionally charge the pieces of clothing displayed and put them in the context of travelling and Italy. The Vespa is the central item that triggers these associations. The picture in the background places the Vespa in an environment that could be Italy. The picture additionally creates the impression that it was taken from a moving vehicle. What is important about this example is that it is the text on the window (I would like to have a Vespa) is needed to identify the vehicle and to activate the pertinent associations.



Aggregates like the one shown below facilitate and accelerate comprehension.



The above system that is frequently found in construction manuals offers a fast and good overview over the material needed. Additionally no translations are needed as the numbers and depictions are understandable without further explanations.

All examples show the ambiguity of images and graphic representations. They accelerate perception but can convey unequivocal meaning only in a very limited way. Images always need a context that supplies a direction for interpretation and/or associations. Exceptions from this rule are those images that can be clearly identified as having only one valid interpretation within a given society. However, even in these cases societal change might soften values and associations. The role and perceptions of religious symbols e.g. today are sometimes used as fashionable accessories and do not betray their user's religiousness.

The above said also holds for the public space and its highly conventionalized structure of topics and discourses. An observer can understand shop windows without any language only because he/she knows the function of shop windows and their connection to shops. Accordingly, linguistic items like invitations or information like "here you can get...", "here xxx is for sale", "come in" can be omitted without causing a loss of information.

10.3 The limitations of written language

The examples from the above chapter have shown that written language is indispensable for interpreting images in multimodal aggregates.

However, written language as part of multimodal aggregates has its limits, too. The most important limitation of written language is that paraverbal information such as emotions, irony or doubt cannot be transmitted. In literature there are always hints, how to understand an utterance:

"It's okay", he laughed

"It's okay", he hissed

"We broke the vase", Paul and Paula confessed in tears

Texts like the ones above are too long and complicated for use in multimodal aggregates. The construction of aggregates aims at ensuring not only easy and fast perception but also memorization. The property of images/graphics to activate associations and emotions performs this share of an aggregate's task.

.



The same holds for wordplays as shown in the example of the car rental company Sixt. Irrespective of how language is transmitted, there is no linguistic signal to mark wordplay. It is only the context of an utterance that can mark it as word game. Other than is the case with the examples given above, in the below advertising a missing image cannot be compensated for with explanations delivered in language.



(Translation: English lawn, meaning: A more English way of going fast)

The spelling of German is correct at first sight. The word "Rasen" (lawn) can be identified as "englischer Rasen" (English lawn). "Englischer Rasen" is a collocation in German used to describe an extremely well kept lawn without any weeds etc. However, in this context this meaning does not make sense. The basis of the word game is "Rasen". Written with a capital letter it is the German noun for "lawn", spelled with a small "r" it is a verb meaning to go or drive fast. Understood this way, the word "Englischer" also changes meaning. In the literal meaning it an adjective (masculine, singular, nominative). However changing the word class of "Rasen" from noun to verb the adjective becomes an adverb and a comparative of "English" meaning "more English". Obviously, there is not only an English way to go fast, but also a more English way that it associated with the car, which is not only an English car but a famous English sports car.

11 ÜBERBLICKSLITERATUR /LITERATURE

- Barthes, Roland (1969). Die Rhetorik des Bildes. In G. Schiwy (Ed.) *Der französische Strukturalismus*. 159-166. Reinbek bei Hamburg: Rowohlt Verlag.
- Bateman, John A. (2008). *Multimodality and genre: a foundation for the systematic analysis of multimodal documents*. Basingstoke: Palgrave Macmillan.
- Bateman, John A. (2014). *Text and image: a critical introduction to the visual/verbal divide*. London: Routledge.
- Boeriis, Morten; Holsanova, Jana (2012). Tracking Visual Segmentation: Connecting Semiotic and Cognitive Perspectives. *Visual Communication* 11(3): 259-281.
- Borscheid, Peter; Wischermann, Clemens eds. (1995). *Bilderwelten des Alltags. Werbung in der Konsumgesellschaft den 19. und 20. Jahrhundert*. Stuttgart: Franz Steiner Verlag.
- Diekmannshenke, Hans Joachim; Klemm, Michael; Stöckl, Hartmut (2011). *Bildlinguistik: Theorien, Methoden, Fallbeispiele*. Berlin: Erich Schmidt Verlag.
- Domke, Christine (2014). *Die Betextung des öffentlichen Raumes: Eine Studie zur Spezifik von Meso-Kommunikation am Beispiel von Bahnhöfen, Innenstädten und Flughäfen*. Heidelberg: Universitätsverlag Winter
- Dunst, A; Laubrock, J. Wildfeuer, Janina (2018). *Empirical Comics Research: Digital, Multimodal, and Cognitive Methods*. New York: Routledge.
- Friedrich Thomas; Schweppenhäuser, Gerhard (2010). *Bildsemiotik. Grundlagen und exemplarische Analysen visueller Kommunikation*. Basel, Boston, Berlin: Birkhäuser.
- Glenberg, Arther; Langston, William (1992). Comprehension of Illustrated Text: Pictures Help to Build Mental Models. *Journal of Memory and Language* 21:129-151.
- Große, Franziska (2011). *Grundbegriffe und Methoden der linguistischen Bildanalyse in Text- und Diskursumgebungen*. Frankfurt: Peter Lang.
- Holmqvist, Kenneth; Wartenberg, Constanze (2005). The role of local design factors for newspaper reading behaviour – an eye-tracking perspective. *Lund University Cognitive Studies* 127:1-21.

- Lefebvre, Henri (1974). La production de l'espace. L'homme et la société. Sociologie de la connaissance marxisme et anthropologie 31-32: 15-32.
- Lobinger, Katharina (2012). Visuelle Kommunikationsforschung: Medienbilder als Herausforderung für die Kommunikations- und Medienwissenschaft. Wiesbaden: Springer.
- Sachs-Hombach, Klaus (1999). Bildgrammatik. Interdisziplinäre Forschungen zur Syntax bildlicher Darstellungsformen. Magdeburg: Scriptorum.
- Schirl, Thomas (2001). Text und Bild in der Werbung: Bedingungen, Wirkungen und Anwendungen bei Anzeigen und Plakaten. Köln: Herbert von Halem.
- Scollon, Ron; Scollon, Suzi (2003). Discourses in Place. Language in the material world. Londong: Routledge.
- Stöckl, Harmut (2004). Die Sprache im Bild – das Bild in der Sprache: zur Verknüpfung von Sprache und Bild im massenmedialen Text. Konzepte, Theorien, Analysemethoden. Berlin: De Gruyter.
- Straßner, Erich (2002). Text-Bild-Kommunikation, Bild-Text-Kommunikation. Tübingen: Niemeyer.
- Spitzmüller, Jürgen (2013). Graphische Variation als soziale Praxis. Eine soziolinguistische Theorie skripturaler „Sichtbarkeit“. Berlin: De Gruyter (
- Warnke, Ingo; Busse Beatrix eds (2014). Place-Making in urbanen Diskursen. Berlin, Boston: Mouton de Gruyter.
- Zantides, Evripides (2014). Semiotics and Visual Communication. Concepts and Practices. Newcastle upon Tyne: Cambridge Scholars.