TexBiG Dataset Document Layout Analysis

Annotations Guideline

Annotation Guideline

General hints:

- a. Please do not use the rectangle feature in CVAT but the polygon feature—even when the region of interest is rectangular.
- b. Draw polygons that cover each object completely—even if this causes areas to overlap.
- c. There are no un-annotated areas besides the background.
- d. The labelling order is from the top left corner to the bottom right.
- 2. Textual layout elements (such as text, heading, captions):
 - a. Textual layout elements are combined if they are separated only by line breaks or paragraph breaks.
 - b. Textual layout elements are separated if they are arranged in a different column.
 - c. For textual boxes the polygons need to cover only the outer boundary. The result should be a convex polygon.
 - d. Empty space resulting from indentations or line breaks should be included into the polygon, if they are not part of the outer boundary.
- 3. Disconnected objects or groups of elements:
 - a. Elements with void areas (such as a frame with large empty space) can be drawn with several bounding boxes / polygons. Use the group feature to combine these elements into a single element.
 - b. The overlap of grouped elements should be as small as possible.

Rule Definition

Content Description: describes what is seen physically in the document image.

[Is] describing the content of the layout element, what can be seen on the document image as part of a class.

[Is not] describes what the layout element is not contentwise.

Class Relations: describes the relations between classes.

[Can include] classes mentioned here can be either fully included in the same content area or overlap, they might refer to different content that overlaps.

[Consumes] if a class is nearby a class that consumes this class, the annotation of the consumed class becomes part of the other class. This means that the region is only part of the consuming class, hence the other class vanishes.

[Have to include] class that can only exists if at least one element of another class can be found within the class.

[At least one] minimum of one, but can also be an arbitrary large number of classes.

[Is grouped if] related content that could be seen as separate class instances, but is actually grouped to a single instance if the condition is fulfilled.

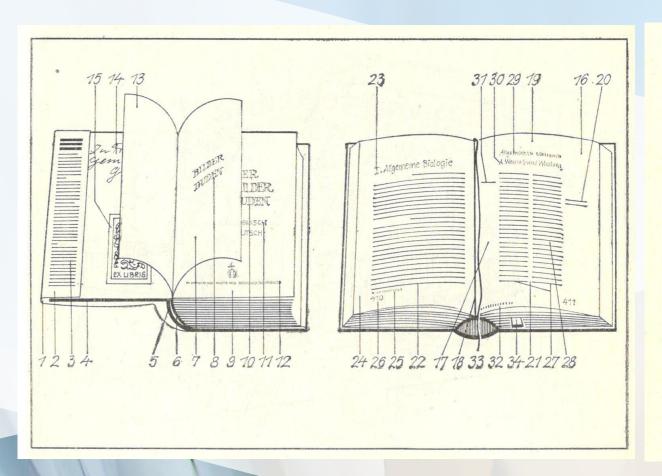
[Is always included in] must be included in another class.

[Cannot be a part] counter condition to [consumes], if a class is consumed it cannot be part of a specific other class since it would always get consumed by that other class.

[Is either one] one of the following descriptions defining the class, only one can be chosen, not multiple.

[Follows] another class must appear somewhere in the document previously, otherwise the class is not valid.

General Structure of a Book



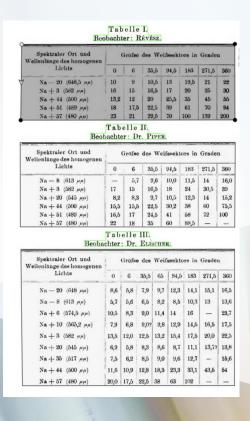
```
Buchumschlag m (Schutzumschlag m)
 2 3 4 5
       Umschlagklappe f
       Klappentext m
       Buchdecke f
       Buchrücken m
       unteres Kapitalband n
       Vortitelseite f
       Schmutztitel m
       Titelblatt n
       Titel m
11
       Untertitel m
12
       Impressum n
13
        Vorsatz m
14
       handgeschriebene Widmung f
15
       Exlibris n
16
       Druckseite f
17
       innerer Papierrand m
18
       Falz m
       oberer Seitenrand m
20
       Vorderkantenrand m
21
       unterer Rand m
22
        Schriftsatzgröße f
23
24
       Kapitelüberschrift f
        Sternchen n
25
       Fußnote f
26
        Seitenzahl f; Pagina f
27
       zweispaltige Satzfläche f
28
        Spalte f
29
        lebender Kolumnentitel m
30
        Untertitel m
31
        Randbemerkung f
32
        Signaturzeichen n
33
        Lesezeichenband n
34
        loses Buchzeichen n
```

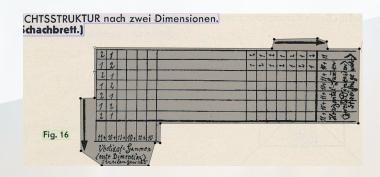
Table

[Is] an ordered arrangement of information or data, typically in rows and columns, or possibly in a more complex structure.

[Has often] a caption.

[Can include] noise.





Image

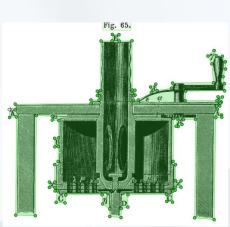
[Is] a two dimensional optical representation of something.

[Can be] a graphic, a photography or a drawing.

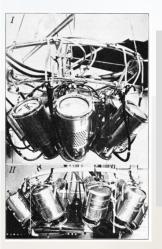
[Can be] a set-up, scheme, curve, technical drawing, graphical recording, musical notation, an object, an interior or exterior view, a person, floor plan, section, map or an illustration.

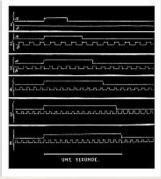
[Is not] a logo, decoration or a frame.

[Can include] noise.









Equation

[Is] a sequence of letters, numbers, formula symbols or words for the abbreviated expression of mathematical or physical facts or relationships.

[Can] have a caption (a reference number, but as a caption class).

[Can include] noise.

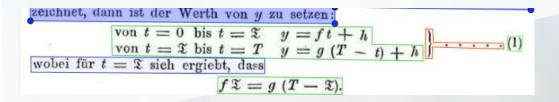
lichem Querschuitt in No. 6 des Centralblattes.

Von einem Abonnenten d. Bl. aufmerksam gemacht, theile ich mit, daß in Formel 19 zwischen den beiden Gliedern der Parenthese + statt — zu setzen ist. Die Formel 20 wird alsdann

$$\tan\left(\beta \operatorname{lognat} \frac{L}{l_o}\right) = 2\beta \frac{\frac{L}{l_o} - 1}{\frac{L}{l_o} + 1}$$

Die Werthe der anschließenden Tabelle ändern sich infolge dessen nur wenig. Die berichtigte Tabelle lautet:

without caption

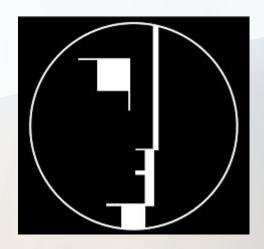


with caption

Logo

[Is] a graphically designed sign of a journal or book etc. It can be a brand or a component of a specific appearance. [Can include] only noise





Editorial Note

[Is] every information that originates from the editor, the publisher or arises from the printing process.

[Is not] noise or a caption.

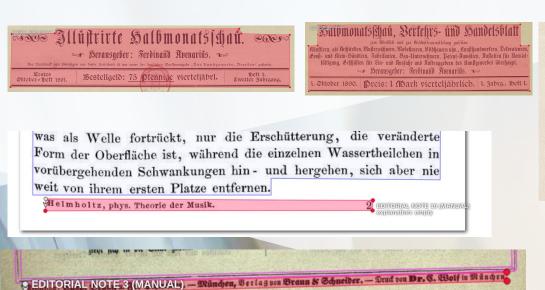
[Is not] an occasion for a footer.

[Examples] "Bogensignaturen", publication details, additional information about the journal.

[Consumes] adjacent decorations.

[Can include] only noise.

explanation: empty



COPYRIGHT 1925 BY ALBERT LANGEN VERLAG
MÜNCHEN

ALLE RECHTE, AUCH DAS DER REPRODUKTION,
VORBEHALTEN
DRUCK HESSE & BECKER • LEIPZIG
KLISCHEES DR. VON LÖBBECKE & CO. • ERFURT
TYPOGRAFIE, UMSCHLAG UND EINBAND:
L. MOHOLY-NAGY • BERLIN

Decoration

[Is] a "artistic" or structuring element.

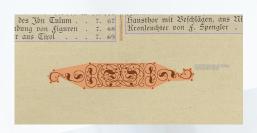
[Is not] a dash for page numbers.

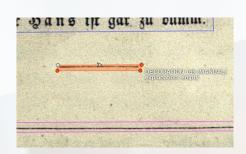
[Examples] an ornament, a frieze, a separator.

[Cannot be a part] of editorial note.

[Is grouped if] elements are directly connected or intersect.

[Can include] noise, text.

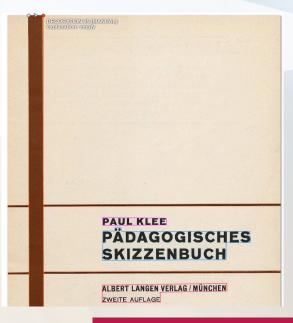




Unsere Farbenschwelle ist also derjenige Betrag eines farbigen Spektral-

In Berichung auf diesen Punkt hat sich die Young-Helmholtzsche Theorie gemäß der Stellung, die sie den Kontrastphänomenen gegenüber einnimmt, nicht als förderlich und anregend erwiesen.

1*

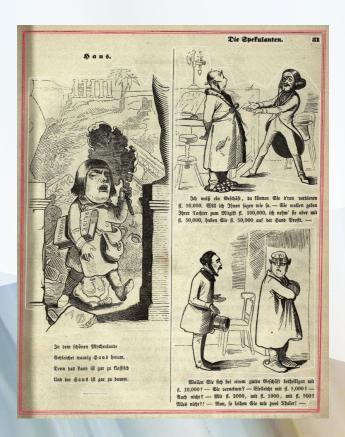


Frame

[Is] the border of an advertisement, image, text or page.

[Is] a single frame or a group of at multiple frames.

[Can include] all other classes.









Advertisement

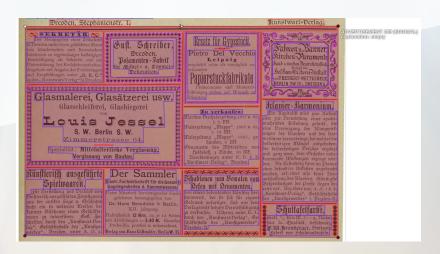
[Is] always rectangular.

[Is] a page-region.

[Can include] decorations, text, images, tables, logos, frames or noise.

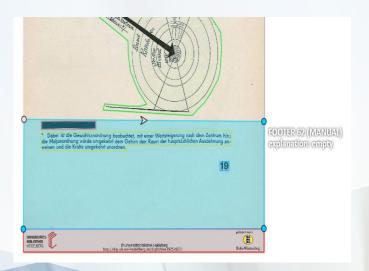


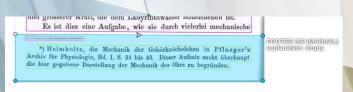


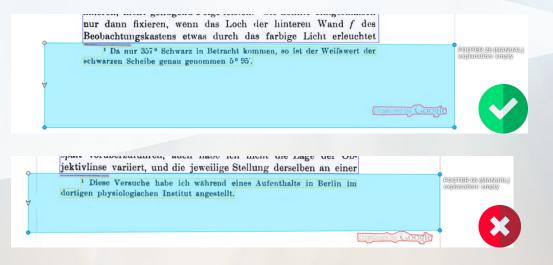


Footer

- [Is] a specific area of a page. The footer follows the main text, or body.
- [Is] used by the author and not by editor or during the printing process.
- [Is] always made up of four points (a trapeze).
- [Is] found only once per page.
- [Is not] found inside a frame, but can overlap with it.
- [Have to include] [at least one] footnotes, page numbers, decoration, author, image.
- [Can include] noise or editorial note (e.g. Bogensignaturen).

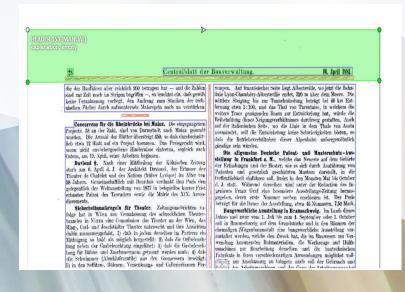


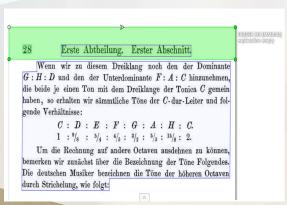




Header

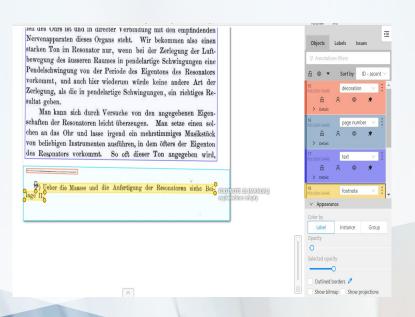
- [Is] detached from the main text at the top (the head) of text pages.
- [ls] used by the author and not by editor or during the printing process.
- [Is] always made up of four points (a trapeze).
- [Is] found only once per page.
- [Is not] a page number or a column title or a heading.
- [Is not] found inside a frame, but can overlap with it.
- [Have to include] [at least one] page number, column title, author, decoration.
- [Can include] noise.





Footnote

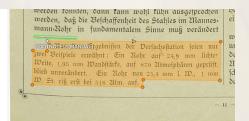
[Is] a note, legend, remark, reference, translation, or further explanation of a word or passage of text by the author. [Is not] part of the main text and [is not] a marginalia. [Can include] images, noise





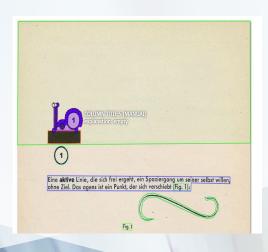
spalt vorüberzuführen, auch habe ich nicht die Lage der Objektivlinse variiert, und die jeweilige Stellung derselben an einer

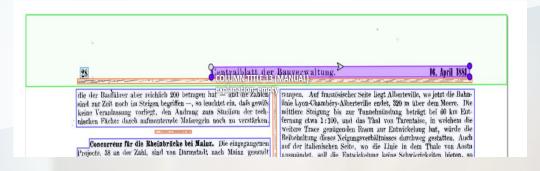
Diese Versuche habe ich während eines Aufenthalts in Berlin im EGOTNOTE 76 (MANUAL) dortigen physiologischen Institut angestellt.



Column Title

[Is] a text usually placed at the top - in the header - of the page or column.[Is not] page number, header or author.[Can include] noise.





5

cotumn ville 74 WANDANgigkeit d. Farbenschwellen von d. achromatischen Erregung.

Page Number

[Is] a consecutive numbering. Page numbers are spatially separated from the column(s).

[Is] usually found on every page OR on no page.

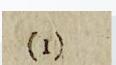
[Is] found only once per page.

[Is not] single numbers from the "Bogensignatur".

[Consumes] adjacent decorations, directly related to the page number.

[Can include] noise.

11 208 PAGE NUMBER 83 (MANUAL) explanation: empty Bauhaus-Universität Weimar



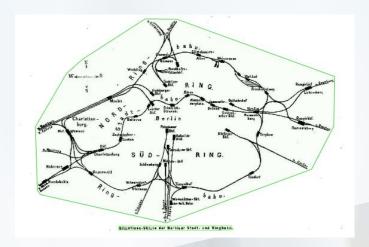
Caption

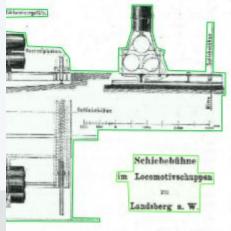
[Is either one]

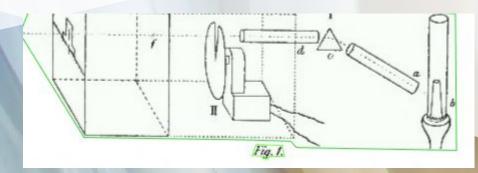
- a. an explanatory text nearby the image/table eventually including a reference number
- b. a reference number nearby the image/table without an explanatory text [Is not] explanations surrounding the reference numbers in the main text.

[Is not] text that is directly part of the image.

[Can include] noise.

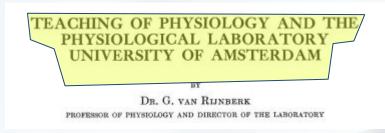


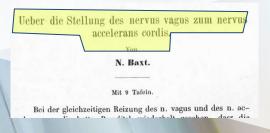




Heading

[Is] name for a book or an article within a journal.[Is not] the header.[Can include] noise.





Aus dem psychologischen Institut der Universität Göttingen.

Ther die Abhängigkeit der Farbenschwellen von der achromatischen Erregung.

Von
Dr. Géza Révész.

§ 1. Einleitung.

Die Young-Helmholtzsche Farbentheorie geht bei der Erklärung der Farbenempfindungen in erster Linie von der Betrachinner des obiektiven Lichtreizes aus ieder Farben-

Subheading

[Is] name for a section or the sub-head of the heading.[Follows] after at least one heading.[Can include] noise.

experiments on mammals they work in groups, and in those on man in couples. Formerly an oral introduction of a quarter of an hour was given before the beginning of every practical work-period. Now the students prepare themselves by consulting the Guide to Physiological Practical Work, a book of 469 pages, written by the director, and amply provided with reproductions of appliances and diagrams.

SCIENTIFIC RESEARCH

With regard to scientific research, a great deal of freedom is allowed the worker in the laboratory. In

achromatischen Erregung.

Von Dr. Géza Révész

1. Einleitung Subheading 41 (MANUAL)

Die Young-Helmholtzsche Farbentenderung der Erklärung der Farbenempfindungen in erster Linie von der Betrachtung des objektiven Lichtreizes aus; jeder Farbenoder Lichtempfindung entspricht ein objektiver Reiz, der aus

Heading

Subheading



Author

[Is] the creator or originator of any written work.
[Is not] the editor, only if the editor is also the author.
[Can include] noise.

[Examples] could be the full name but also in the form of abbreviations with and without a point. The author can be above or below the title or at the end of the text and can also be named in the header but also anywhere else on the page.

Index hat mneen gehört zu den perente ununterbrochene Nahrungs. n. für das nördliche Klima * Kün rüne Abarten der Rhamneen alten Beigi tdeckung dürfte fich zu einer Zeit fein. ung" der Seideninduftrie gedaß es nur ortaetmannägli R. Liiders. Auf it Deutschland (Braunschweig, öligen und Dechelde) immer mehr hinter höchste fäh





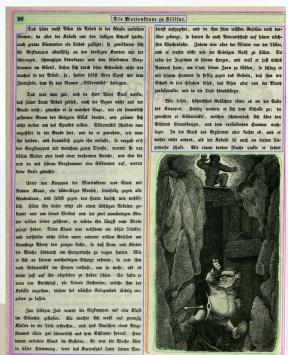




Text

[Is] a thematically and / or functionally oriented, coherent linguistic or linguistic-figurative complex that was created with a certain communication intention. Text can be divided into columns.
[Can include] noise, captions or dropped capital.







Noise

[Is] an element subsequently added to the published work.

[Is] a significant distraction / area on the page (e.g. not a single inc dot).

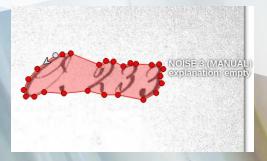
[Examples] are stamps, handwritten notes, comments, markings, digitalization note etc.

[Consumes] all elements that are semantically part of the noise e.g. text inside the stamp, content of the handwritten note, logo in the digitalization note.











Example of textual layout elements with convex boundary







Salbmonatskau, Berkehrs- und Sandelsblatt zum überblick und zur Geschäftsvermittelung zwischen Künstlern, als Architesten, Musterzeichnern, Modelleuren, Sidhauern usw., Kunsthandwersern, Desorateuren, Groß- und Klein-Händlern, Fabrisanten, Bau-Unternehmern, Patent-Anwätten, Anstalten für Bervielfältigung, Geschäften für Sin- und Aussuhr und Auftraggebern des Kunstgewerbes überhaupt. Serausgeber: Ferdinand Avenarius. 1. Oktober 1890. Preis: 1 Mark viertelsährlich. 1. Jahrg., Ibest 1.

Textual elements, connected and disconnected





ber Kaften war aufgesprungen, Melchior hatte ben ersten gescheideten Streich in seinem Leben gemacht und kann Schreiber dieses am besten dem günftigen Leser jagen, daß der Melchior ber glücklichste Gatte und ber fröhlichste Berückenmacher unter der Sonne ist, benn ich bin selbst der Melchior Unter der Aunior und habe meine

Lebens- und Liebesgeschichte zur Aufmunterung für alle bie verfaßt, welche, wenn sie noch so viel sind und verstehen, doch nicht meinen, heirathen zu können, wenn sie kein Geld und kein Umt haben. Auch in unsern ausgeklärten Zeiten hat jeder solch eine spukhafte Geldtiste, die sich ihm öffnet, wenn er durch die That zeigt, daß er ein rechter Kerl ist. Dieses ist die Moral von der Sache. — Mit freundlichem Gruße und mit dem



ber Kaften war aufgesprungen, Melchior hatte den ersten gescheidten Streich in seinem Leben gemacht und kann Schreiber dieses am besten dem günftigen Leser sagen, daß der Melchior der glücklichste Gatte und der fröhlichste Berückenmacher unter der Sonne ist, denn ich bin selbst der Melchior Aunior und habe meine

Lebens- und Liebesgeschichte zur Aufmunterung für alle bie versaßt, welche, wenn sie noch so viel sind und verstehen, doch nicht meinen, heirathen zu können, wenn sie kein Geld und kein Amt haben. Auch in unsern ausgeklärten Zeiten hat jeder solch eine spuckhafte Geldtiste, die sich ihm öffnet, wenn er durch die That zeigt, daß er ein rechter Kerl ist. Dieses ist die Moral von der Sache. — Mit freundlichem Gruße und mit dem

Textual elements, connected and disconnected



geredet hatte von dem ungeheuren Andegalde der Saby Rothingnir, reifte der Lord nach Karlsan wim Myladon nach Baben; fie follen eine febr glidfliche We gerführt, umb fich vor dem Tade des Seords einmal in Petersburg, und einmal in Petersburg, und einmal in Petersburg, und einmal in Vederfen haben.

Sah Rarr! lachte Junior, bas war nur ein Gefpag mit ber Lordicait, ich bin nichts und habe nichts, und will bich bennoch nachsten Sonnag heirathen.

Bebt ichmedt mir tein Biffen mehr, fiohnte Babet umb legte erbleichend Meffer umd Sabel nieder. Du leichtfinniger Burjde bu! Bas ift benir aber in der großen eifernen Rifte hinter bir ? —

3ft das dumme Ding, brummte Junior, mir auch hierber nachspagiert! In dem Kaften ift viel Geld, aber ich tann nicht bagu.

Wemm wir tein Gelo und tein Amt haben, rief Bab et, so tonnen wir uns anch nicht heirathen. D! ich unglidtliche Madden, meinen guten Tienft habe ich anfegeben und einen Mann trieg ich nur boch nicht! babet liefen ihr bie hellen Ebranen bie Baden herunter.

Das ift langweilig, fing auch Junior en zu weinen, doß wan nicht hetratfen kann, wenn man kein Geld har; jetzt Jabe ich endlich gelerut, mos es heitst fiel jangweilen. Meer ad wost jetter er, plostich sich sollend, hinzu, und umarmte seine Brant, wenn wir auch lein Geld haben, ich verstehe mich ja auf? Sanatristien und ver ert recht, do heirathe ich bich doch mb wir werden zusummen ein Veridenmacher.

Auf einmal knallte es hinter ihnen wie ein Biftolenichus, daß sie erschrocken von einander fuhren. D! Freude,



geredet hatte von dem ungeneren Nabelgelde der Lab. Robbling nir, reifte der Lord nach Karls ab um Malady mach Baden; Baden; fie follen eine jehr glättliche Ete ger führt, und ich vor dem Zode des Lords einmal in Paris, einmal in Peter Burg, und einmal in Papel, das lette Mal sonar und ein ange Etnte wieder, geffen hoben.

Au nior umd Bade 1 hotten lich daggene balt ihre Liebe ertfart und durch tausiend Külfe bewiefen. Dam schiften lie ber Gräftn mit Zunt und Grug ihren Bagen wieder gurüd, und befeldlen lich in dem Wirtsphaue einen Berlohunghschmauß, da ut ihn der Berloh, in der Geschwinshielle bereich schoffen tonnte. Alle sie den wierten Gang, einen vortrefflichen Lumurofibraten verzehrt hatten und auf dem Reiguberdin gometteen, sagte endlich die oreflundige Babet! Aber wenn du ein so vorressmer und reisser derer bift, fannst du und wilfit du mich auch geitruffen?

Sah Narr! lachte Junior, bas war nur ein Gespaß mit ber Lorbichaft, ich bin nichts und habe nichts, und will bich bennoch nächten Sonnag heirathen.

Best ichmedt mir fein Biffen mehr, fiohnte Babet und legte erbleichen Meffer und Gabel nieder. Du leichtfinniger Bursche bu! Bas ift benn aber in der großen eifernen Rifte hinter bir ? —

Ift das dumme Ting, brummte Junior, mir auch hierher nachspaziert! In dem Kasten ist viel Geld, aber ich kann nicht basu.

Menn wir tein Gelo und tein Ant haben, rief Ba b c t, fo tomen wir und and nicht heirathen. D! ich unglieftliche Mächen, meinen guten Dienft habe ich aufgegeben und einen Mann trieg ich nun bach nicht! babei fiefen ihr bie hellen Trongen die Barten beruntet.

Das ift langweilig, fing auch Junt or an zu weinen, dom nicht hetrafpen tann, wenn man tein Gelb hat; jest habe ich andich gelernt, was de heith find langeweilen. Ber ach was! feste er, plöstlich sich fassen, ich vertiebe mich ja auf's Daartritten und tein Gelb haben, ich vertiebe mich ja auf's Daartritten und be u erit recht, de heitrafte ich bich doch und mir werden zusammen ein Versichennacher.

Auf einmal knallte es hinter ihnen wie ein Biftolenichus, daß sie erschrocken von einander fuhren. D! Freude

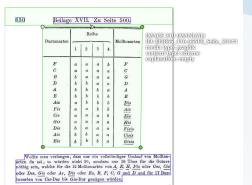
Image included in text or not?

560 Dritte Abtheilung. Achtzehnter Abschnitt,

Bewegung in die Consonanz einer Octave oler Quinte (Duodecime)
übergeben sollen. Sie sollen vielmehr in eine solche Consonanz nur
in Gegenbewegung treten, die eine sinkvad, die andere steigend.
Dasselbe wirde im zwickimmigen Satze für den Kinkhang gelten.
Der Sim diese Gesterte ist wohl nur für, dass jedeman, wo die
per Sim diese Gesterte in der Sinkhang den die sinken der Sinkhang gelten,
sen, als einen relativen Robenatand gegen einander errammende,
siehe soffen der Sinkhander sollten siehe sollten siehe sollten siehe sollten siehe sen,
sen die gegen gestelne siehen im der Sinkhander siehen siehe sie

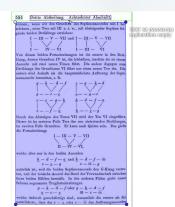


Eine andere Regel der Stimmführung, betreffend (Onlightistyne miss end nochtion sen unharmonischen Quarstand, ergab sieh wolkcyfildhildhild (miph) dem Bedirinis der Singer. Was aber fin den Singer selwer zu tereffen ist, muse natürlich auch dem Hörer innere als ein ungewöhnlicher und gewungener Schaltt eracheinen. Unter Querstand versiehen under Singer selwer und dem Singer selwer selwer und dem Singer selwer und d



Images separate from text

- contain decorative or structuring non-typographic elements (e.g. frames, borders)
- are thematically separate from the flow of text



ung emereten mussen.

B. Klänge, welche neben dem Grundton die Duodeeime hören lassen. Ein Beispiel solcher Klänge geben die engen gedackten Pfeifen der Orgel (Register, Quintaten). Diese verhalten sich ebenso wie die, welche bloss Octaven als Begleitung des Grundtones hören lassen.

 Quinte. Grundtöne 200 und 301 mit Obertönen 600 und 903. Erster Differenzton;

> 903 - 600 = 303Quinte = 301

Zahl der Schwebungen: 2.

b. Quarte. Grundtöne: 300 und 401, mit Obertönen 900 und 1203. Erster Differenzton:

> 1203 - 900 = 303Grundton = 300

Zahl der Schwebungen: 3.

Schwebungen der Terzen können auch in diesem Falle nur durch die schwachen zweiten Differenztöne eintreten.

C. Klänge, bei denen neben den Grundtönen gleichzeitig die Octaven und Duodecimen als Obertöne hörbar und 1203 der beiden primären Töne erzeugt worden wären, wenn diese existirten. Diese Schwebungen der Quarte sind nun schon sehr schwach, auch bei starken primären Tönen. Zu ihrer Beobachtung muss man ganz ungestört sein, und grosse Aufmerksamkeit anwenden.

Kaum noch wahrzunehmen, auch unter den günstigsten Bedingungen, sind die Schwebungen der unreinen grossen Terz. Nehmen wir die Schwingungszahlen der primären Töne 400 und 501, so ist 501 — 400 = 101 Combinationston erster Ordnung

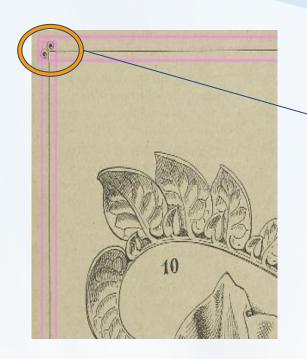
Die Töne 202 und 198 geben 4 Schwebungen. Scheibler hat diese Schwebungen der uureinen grossen Terz noch gezählt, ich selbst habe sie unter günstigsten Bedingungen auch wohl noch zu hören geglaubt, aber jedenfalls sind sie so schwer wahrzunehmen, dass sie bei der Bestimmung des Unterschiedes von Consonanzen und Dissonanzen nicht mehr in das Gewicht fällen können.

Daraus folgt also, dass die verschiedenen Intervalle, die der Terz benachbart sind, durch den Zusammenklang zweier einfachen Töne gleichmässig hergestellt werden können, ohne dass ein Unterschied des Webblewege stattfüg verso die sicht einzestelt des

"images"/schemes in text

- contain only typographic elements
- are meant to be read with the flow of text

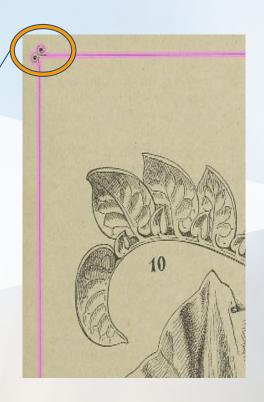
Grouped element overlap and void











Acknowledgement

This project would not have been taken shape without our team: Sascha Breithaupt, Johannes Hess, Josephine Tiede and Ina Tscherner.