

# Documents and Texts in TEI XML: Representing Wittgenstein's Text Alternatives

Alois Pichler

*Wittgenstein-Schlick-Scholz workshop*

Bergen 11.6.2024

# Program

- 1) Two basic distinctions:
  - Writing acts that are text-constitutive vs. Writing acts that are not
  - Documents vs. Texts
- 2) Encoding simple cases of Wittgenstein text alternatives
- 3) Encoding cases of Wittgenstein text alternatives that require to separate encoding of document and text levels

I HATE

I CAR/S

# Writing acts that are text-constitutive vs. writing acts that are not

- Deletion of “R” and insertion of “T” in “CARS” produces a new text, namely “I HATE CATS”
  - “I HATE CATS” is text alternative to “I HATE CARS”
- Deletion of “R” and insertion of “T” in “HARE” does not produce a new text but is an orthographic correction of “HARE” to “HATE”
  - “I HARE CARS” is not text alternative to “I HATE CARS”
  - “I HARE CATS” is not text alternative to “I HATE CATS”
- The same **document** → Different **texts**

# Different documents → The same text

	Document level	Text level
I love <sup>cats</sup> <del>cars</del> .	I love	cats <del>cars</del> .
I love <sup>cats</sup> <del>cars</del> .	I love	cats. <del>cars</del> .
I love <sup>t</sup> <del>cars</del> .	I love	t <del>cars</del> .
I love <del>cars</del> cats.	I love	<del>cars</del> cats.
I love <del>cars</del> //cats//.	I love	<del>cars</del> //cats//.
I love <del>cars</del> <sup>t</sup> cats.	I love	ca[r t]s.
I love <del>cars</del> carts.	I love	ca <del>r</del> ts.

I love **cars**.

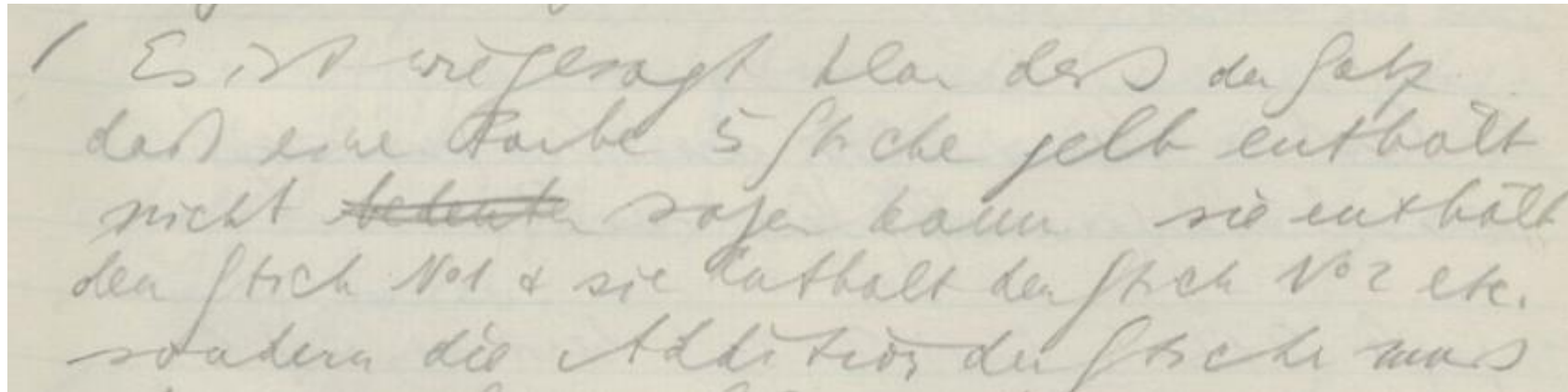
vs.

I love **cats**.

# Some examples of text alternatives from Wittgenstein's Nachlass, and WAB's encoding of them

Examples of cases where the encoding of the document level and the encoding of the text level go straightforwardly together

Ms-106,95[5]et97[1] «bedeuten» is replaced by «sagen» (Sofortkorrektur)



```
<choice type="dsl"><orig type="alt1"><del type="d">bedeuten</del></orig>  
<orig type="alt2">sagen</orig></choice>
```

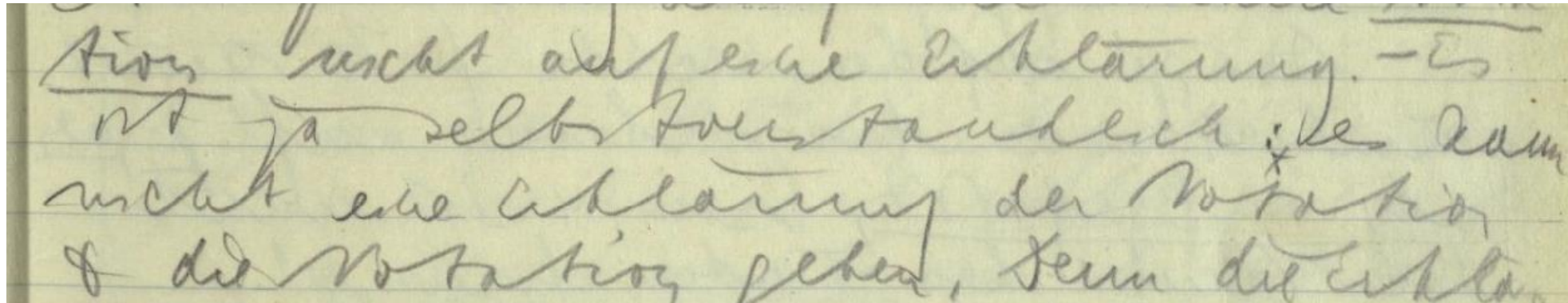


Ms-114,104v[2] «ihm» is replaced by «dem Befehl» (Sofortkorrektur)



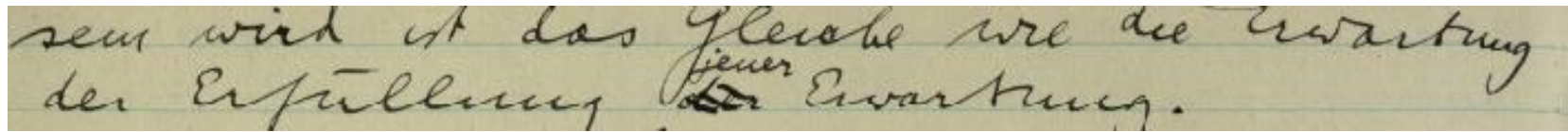
```
<choice type="dsl"><orig type="alt1"><del type="d">ihm</del>  
<lb/></orig> <orig type="alt2">dem Befehl</orig></choice>
```

# Ms-109,1[3] text alternative is decided in favour of colon rather than comma



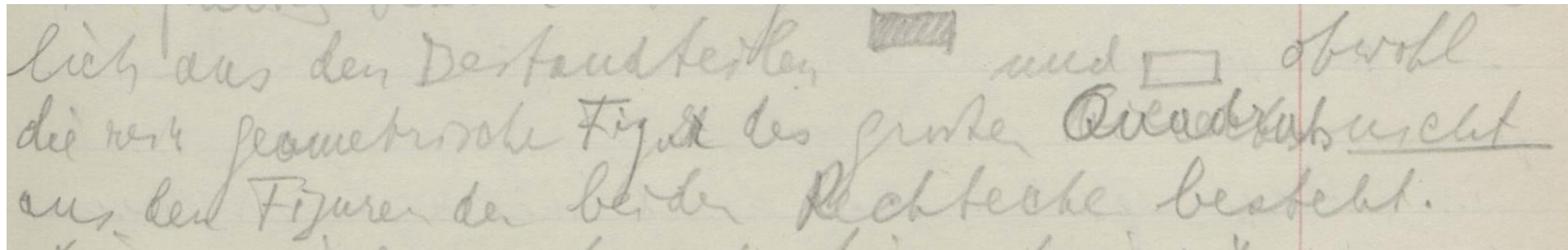
```
<choice type="dsf"><orig type="alt1">:</orig>  
<orig type="alt2"><del type="d"><add  
rend="el">,</add></del></orig></choice>
```

Ms-114,103r[2]et103v[1] «der» is replaced by «jener»



```
<choice type="dsl"><orig type="alt1"><del type="d">der</del></orig>  
<orig type="alt2"><add rend="i">jener</add></orig></choice>
```

Ms-105,43[1]et45[1] «Vierecks» is replaced by «Quadrats»



<choice type="os"><orig type="os1">Vierecks</orig>  
<orig type="os2">Quadrats</orig></choice>

Cases that require to **separate**  
encoding of text level and  
encoding of document level

# Diplomatic representation (→ document level) vs. Normalized / linear representation (→ text level)

	Document level	Text level
	I love	cats
	I love	<del>cars</del> .
	I love	cats.
	I love	<del>cars</del> .
	I love	t
	I love	cars.
	I love	<del>cars</del> cats.
	I love	<del>cars</del> //cats//.
	I love	ca[r t]s.
	I love	ca <del>r</del> ts.

I love **cars**.

vs.

I love **cats**.

# MECS-WIT, A Registration Standard for the Wittgenstein Archives at the University of Bergen:

em [em/2\... | ...]

This code should be applied in **cases where parts of text have to be repeated or extended**, normally as elements in substitutions, in order to comply to the rules stated in the Encoding Guidelines on emendation of cotextually bound text elements, and on encoding below word-level. The first element should contain an ad verbatim transcription, the second element should contain the readings in a simplified form.

May NOT be used recursively

Not to be used in emendations

First element excluded from alpha text

First element excluded from beta text

# Diplomatic representation (→ document level) vs. Normalized / linear representation (→ text level)

Document	Encoding: I love ...	Text
love <sup>cats</sup> <del>cars</del> .	... <choice type="dsl"><orig type="alt1"><del type="d">cars</del></orig> <orig type="alt2"><add rend="i">cats</add></orig></choice>.	I love cars    cats.
love <sup>cats</sup> <del>cars</del> .	... <choice type="dsl"><orig type="alt1"><del type="d">cars.</del></orig> <orig type="alt2"><add rend="i">cats.</add></orig></choice>	
love <sup>t</sup> cars.	... <choice type="em"><orig type="em1">ca<del type="d">r</del> <add rend="i">t</add>s</orig> <orig type="em2"><choice type="dsl"><orig type="alt1">cars</orig> <orig type="alt2">cats</orig></choice></orig></choice>.	
love <del>cars</del> cats.	... <choice type="dsl"><orig type="alt1"><del type="d">cars</del></orig> <orig type="alt2">cats</orig></choice>.	
love cars // cats //.	... <choice type="dsl"><orig type="alt1"><del type="d">cars</del></orig> <orig type="alt2">  cats  </orig></choice>.	
love cars.	... <choice type="em"><orig type="em1">ca<choice type="o_h"><orig type="o1">r</orig><orig type="o2"> t</orig></choice>s</orig> <orig type="em2"><choice type="dsl"><orig type="alt1">cars</orig> <orig type="alt2">cats</orig></choice></orig></choice>.	
love cars.	... <choice type="em"><orig type="em1">ca<del type="d">r</del>ts</orig> <orig type="em2"><choice type="dsl"><orig type="alt1">cars</orig> <orig type="alt2">cats</orig></choice></orig></choice>.	



# Huitfeldt 2006:

The criteria for ensuring that a transcription is suited for normalised reproduction were less easy to formulate, and were dealt with by means of a much more elaborate and formal approach, to be described in the next two sections.

## Formalisation and operationalisation

Consider the following hypothetical example, which is a bit more complicated than the previous one:

Again, an obvious requirement for a normalised reproduction is that orthographic errors are corrected. In this example one will have to mark up the misspelling 'Vather' so that it can be rendered correctly as 'Vater' in the normalised version. Admittedly, orthographic rules are not always clear, and texts are frequently written according to idiosyncratic or inconsistent orthographies. Further complications are that orthography variation is quite often a literary means of expression, and that orthography may in itself be an object of study. In electronic texts spelling affects not only readability, but also retrievability. Therefore, standardisation is much more important in electronic than in traditional editions. While there may be a need to retain the original orthography, as in diplomatic transcription, there is also a need to standardise orthography to some set of uniform spelling rules.

But what about the variant readings? In the current example, 'ein' seems to have been substituted for 'eine', 'großes' for 'große' and 'Haus' for 'Hütte'. At least the following readings seem to be possible:

1. Mein Vater hat eine sehr große Hütte
2. Mein Vater hat ein sehr großes Haus
3. Mein Vater hat ein sehr großes weißes Haus

On the assumption that the inserted 'weißes' is intended to apply to both a) and b), and that the author simply forgot that the insertion of the adjective 'weißes' in a) would require the inflected form 'weiße', perhaps even this is a possible reading:

1. Mein Vater hat eine sehr große weiße Hütte

As with the previous example, some kind of guidance is needed as to whether all or only some of these are to count as possible readings. However, not every possible combination of substituentia makes sense. For example, even though 'Hütte' has not been deleted, 'Mein Vater hat ein sehr großes Hütte' is ungrammatical and not a possible reading. By mechanically selecting one out of every pair of substituentia, one can create a large number of obviously invalid readings. One needs to mark the text up in a way which does not include these.

We started out by making one basic decision: If among two transcriptions interpretational considerations do not decide clearly in favour of the one rather than the other, we would decide in favour of the one which came closest to an ideal of what we called a well-formed text, which we defined as follows:

A well-formed text is a unilinear sequence of orthographically acceptable and grammatically well-formed sentences which together form a coherent text unit.

Our next step was to define what we called an alpha-text:

An alpha-text is as a set of strings derived from a transcription according to a language-specific procedure.

There was one such procedure for each language present in the transcription. The procedures consisted in assigning functions such as inclusion, exclusion and case change to every element type.

Our final step was to define what we called beta-texts. We observed that whereas a well-formed text was defined as a unilinear sequence of sentences, a manuscript with alternative readings was multi-linear rather than unilinear. We defined a transcription's beta-texts (roughly) as follows:

The beta-texts of a transcription is a set of texts derived by excluding, for every substitution, all except one reading, and repeating the process until all possible combinations of readings of the entire text has been exhausted.

Thus, our basic requirement that transcriptions should be well-formed could be reformulated as the requirements that:

- T1 Every alphatext derived from the transcription should be a list of orthographically acceptable graphwords in the language of that alphatext.
- T2 Every betatext derivable from the transcription should be a well-formed text.

To avoid a possible misunderstanding: The set of betatexts derivable from one and the same transcription do not necessarily represent different interpretations of the manuscript in question. If the transcription derives more than one betatext, then that may just as well be regarded as a characteristic of the interpretation in question.

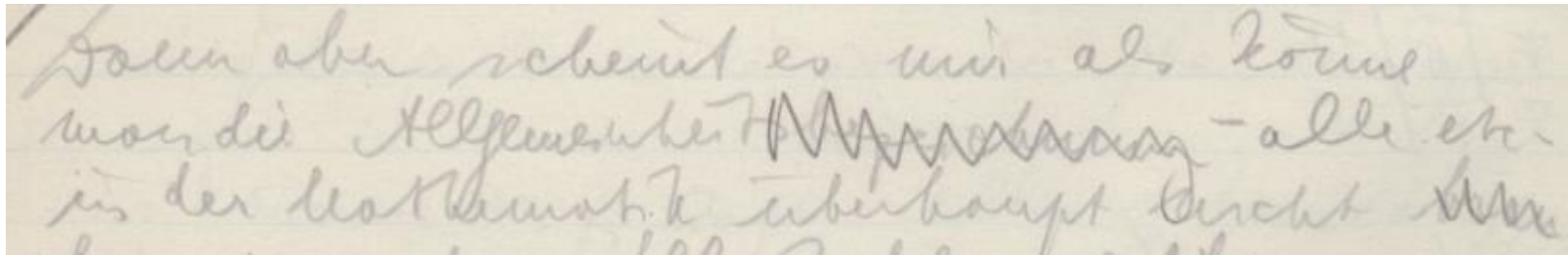
## Transcription Method

So far, we had formulated criteria which allowed us to identify some rather than others among possible transcriptions as acceptable. What remained to be done was to prescribe a procedure which helped the transcriber satisfy these criteria. To this end, we made use of a somewhat indirect strategy: We prescribed a procedure which, given the nature of the manuscripts we were dealing with, was in most cases almost certain to produce a transcription which would not satisfy the criteria. This basic procedure was:

RULE A:

Manuscripts should be transcribed page by page from the front towards the back page, each page vertically line by line from top to bottom, each line horizontally letter by letter from left to right.

Ms-106,90[4]et92[1] «Allgemeinheitsbezeichnung» is replaced by «Allgemeinheit»



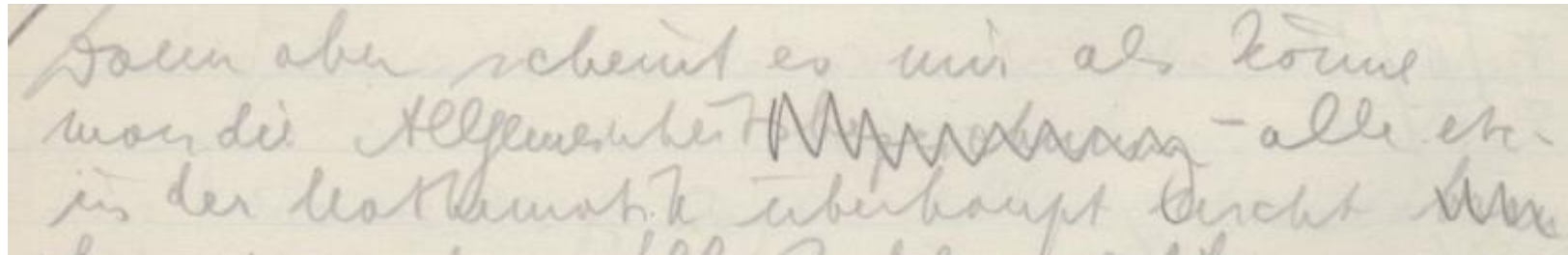
(1) `<choice type="dsl"><orig type="alt1">Allgemeinheitsbezeichnung</orig>  
<orig type="alt2">Allgemeinheit</orig></choice>`

Does not permit to reproduce the document (diplomatic = document transcription!)

(2) `Allgemeinheit<del type="d">sbezeichnung</del>`

Does not permit to extract the text alternatives (linear / normalized = text transcription!)

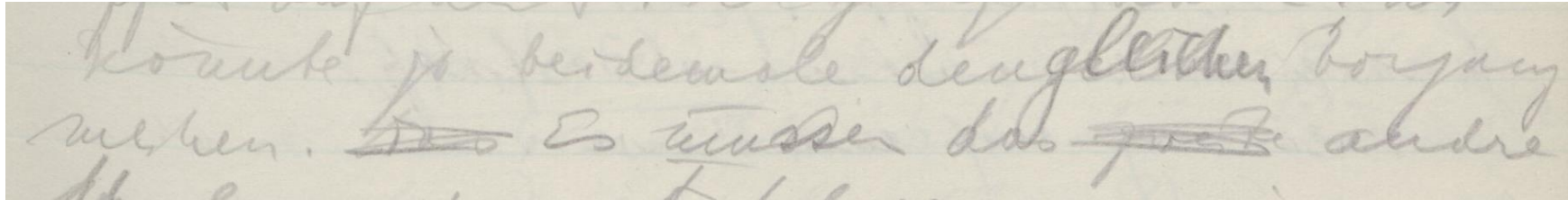
Ms-106,90[4]et92[1] «Allgemeinheitsbezeichnung» is replaced by «Allgemeinheit»



```
<choice type="em"><orig type="em1">Allgemeinheit<del  
type="d">sbezeichnung</del></orig>
```

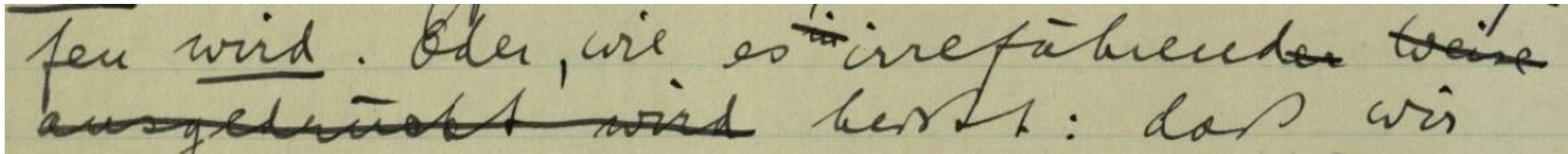
```
<orig type="em2"><choice type="dsl"><orig  
type="alt1">Allgemeinheitsbezeichnung</orig> <orig  
type="alt2">Allgemeinheit</orig></choice></orig></choice>
```

Ms-106,130[5]et132[1] «selben» is replaced by «gleichen» / «denselben» is replaced by «den gleichen»



- den <choice type="os"><orig type="os1">selben</orig>  
<orig type="os2">gleichen</orig></choice>
- <choice type="em"><orig type="em1">den<choice type="o"><orig  
type="os1">selben</orig><orig type="o2">gleichen</orig></choice></orig>  
<orig type="em2"><choice type="dsl"><orig type="alt1">denselben</orig> <orig  
type="alt2">den gleichen</orig></choice></orig></choice>

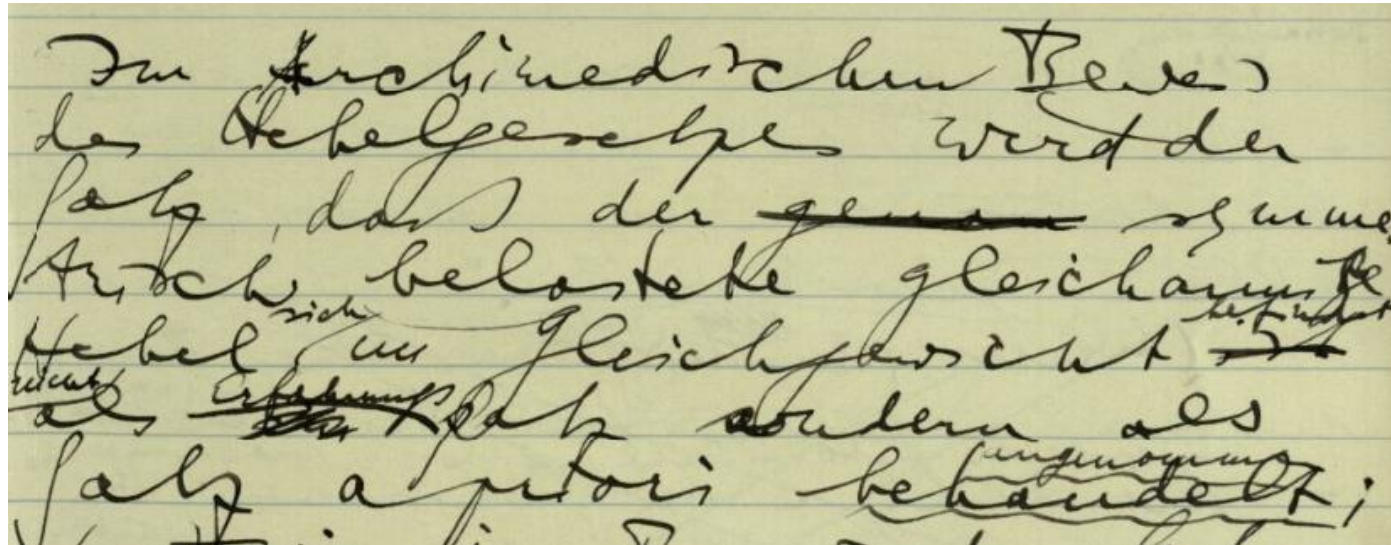
Ms-114,98v[2] «irreführend» is (considered) replaced by «in irreführender Weise» which is replaced by «irreführend»



```
<choice type="em"><orig type="em1"><del type="d"><add rend="i">in</add>  
</del> irreführend<del type="d">er Weise</del><lb/></orig>
```

```
<orig type="em2"><choice type="dsl"><orig type="alt1">irreführend</orig>  
<orig type="alt2">in irreführender Weise</orig> <orig  
type="alt3">irreführend</orig></choice></orig></choice>
```

# Ms-122,25r[3]et25v[1] «Satz» is replaced by «Erfahrungssatz»

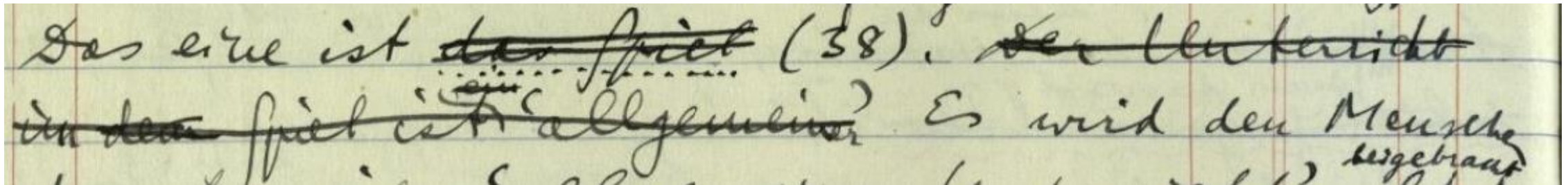


<choice type="em"><orig type="em1"><add rend="im">Erfahrung<add rend="im">Erfahrung</add><choice type="o"><orig type="o1">S</orig><orig type="o2">s</orig></choice>atz</orig>

<orig type="em2"><choice type="dsl"><orig type="alt1">Satz</orig> <orig type="alt2">Erfahrungssatz</orig></choice></orig></choice>



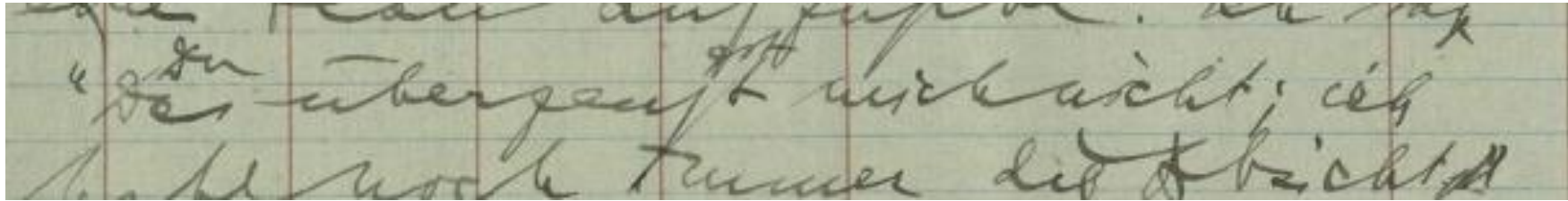
Ms-115,155[2] «allgemein» is replaced by «ein allgemeiner»



```
<choice type="em"><orig type="em1"><add rend="im">ein</add>  
'allgemein<add rend="el">er</add>'</orig>
```

```
<orig type="em2"><choice type="dsl"><orig type="alt1">'allgemein'</orig>  
<orig type="alt2">ein 'allgemeiner'</orig></choice></orig></choice>
```

Ms-128,27[2] «Das überzeugt» is (considered) replaced by «Du überzeugst»

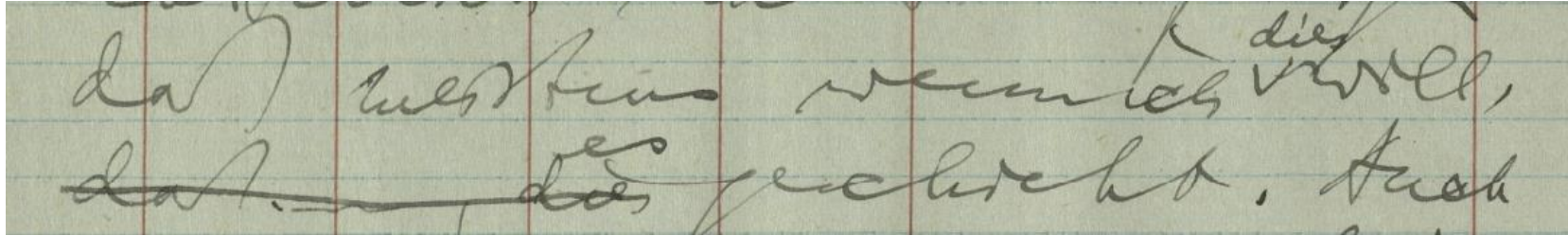


`<choice type="em"><orig type="em1">Das <add rend="i">Du</add>  
überzeugt<add rend="i">gst</add></orig>`

`<orig type="em2"><choice type="s"><orig type="alt1"><c type="c">D</c>as  
überzeugt</orig> <orig type="alt2">Du  
überzeugst</orig></choice></orig></choice>`



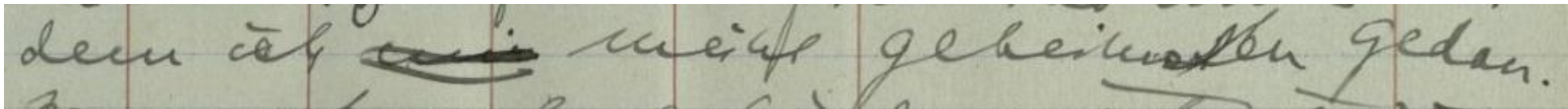
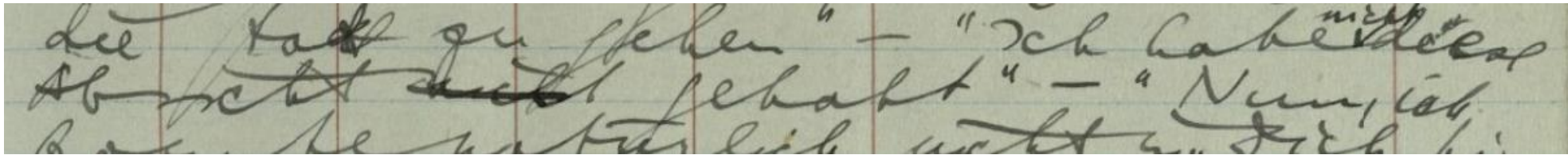
Ms-128,31[5]et32[1] «will, daß ..., dies geschieht» is replaced by «dies will, es geschieht»



```
<choice type="em"><orig type="em1"><add rend="im">dies</add> will, <del type="d">daß ..., dies</del> <add rend="i">es</add></orig>
```

```
<orig type="em2"><choice type="dsl"><orig type="alt1">will, daß ..., <choice type="s"><orig type="alt1">dies</orig>  
<orig type="alt2"><add rend="i">es</add></orig></choice></orig>  
<orig type="alt2"><add rend="im">dies</add> will, <add rend="i">es</add></orig></choice></orig></choice> geschieht
```

Ms-128,27[2] «diese Absicht nicht» is replaced by «nicht diese Absicht» etc.



```
<choice type="em"><orig type="em1"><add rend="im">nicht</add> diese Absicht <del type="d">nicht</del></orig>
```

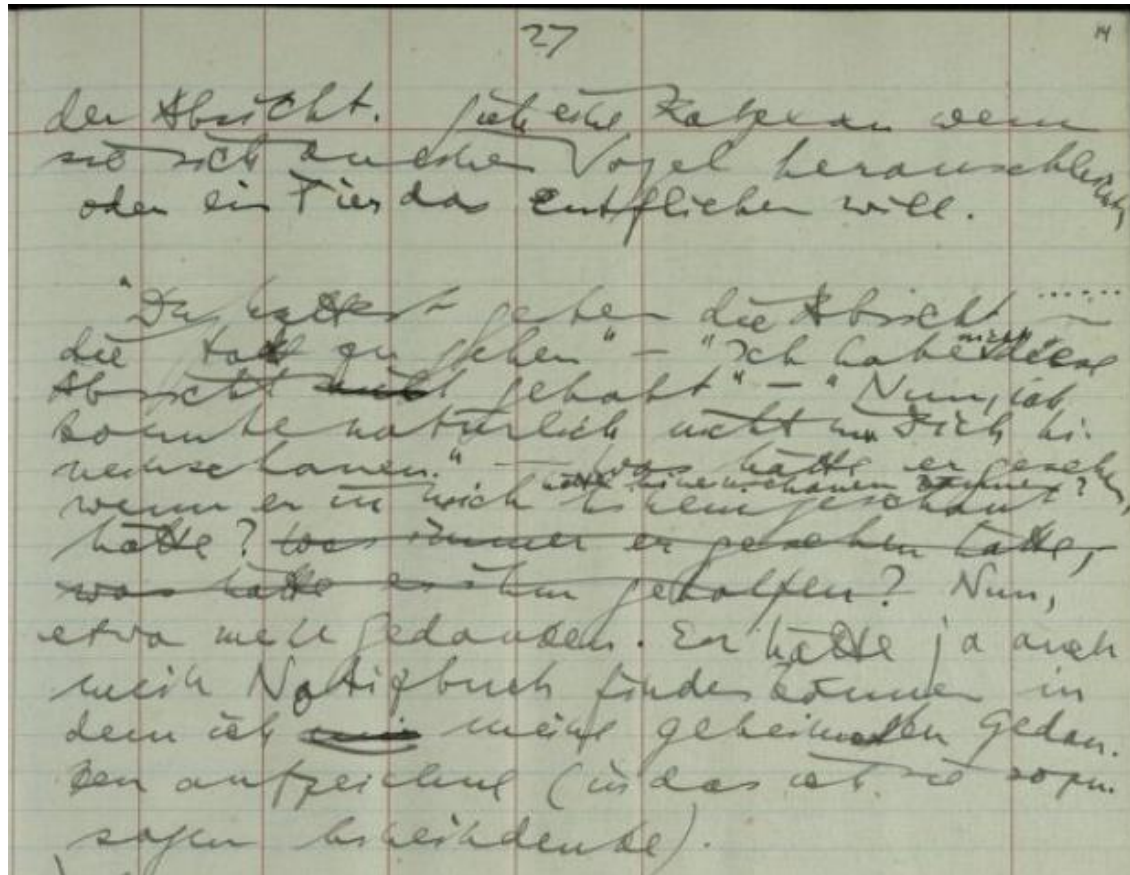
```
<orig type="em2"><choice type="dsl"><orig type="alt1">diese Absicht nicht</orig> <orig type="alt2"><add rend="im">nicht</add> diese Absicht</orig></choice></orig></choice>
```

...

```
<choice type="em"><orig type="em1">geheim<del type="d">st</del>en</orig>
```

```
<orig type="em2"><choice type="dsl"><orig type="alt1">geheimsten</orig>  
<orig type="alt2">geheimen</orig></choice></orig></choice>
```

# Ms-128,27[2] «diese Absicht nicht» is replaced by «nicht diese Absicht» etc.



27

der Absicht. Sieh eine Katze an wenn sie sich an einen Vogel heranschleicht, oder ein Tier das entfliehen will.

"Du hattest gestern die Absicht in die Stadt zu gehen" – "Ich habe <sup>nicht</sup> diese Absicht nicht gehabt" – "Nun, ich konnte natürlich nicht in Dich hineinschauen." – Was hätte er gesehen, wenn er hätte hineinschauen können?

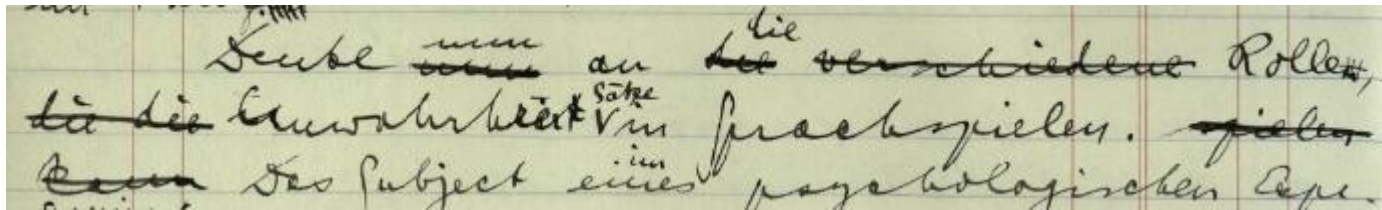
in mich hineingeschaut hätte? Was immer er gesehen hätte, was hätte es ihm geholfen? Nun, etwa meine Gedanken. Er hätte ja auch mein Notizbuch finden können in dem ich ~~mir~~ meine geheimsten Gedanken aufzeichne (in das ich sie sozusagen hineindeute).

27

der Absicht. Sieh eine Katze an wenn sie sich an einen Vogel heranschleicht, oder ein Tier das entfliehen will.

"Du hattest gestern die Absicht in die Stadt zu gehen." – "Ich habe diese Absicht nicht || nicht diese Absicht gehabt." – "Nun, ich konnte natürlich nicht in Dich hineinschauen." – Was hätte er gesehen, wenn er in mich hineingeschaut hätte? || hätte hineinschauen können? Was immer er gesehen hätte, was hätte es ihm geholfen? Nun, etwa meine Gedanken. Er hätte ja auch mein Notizbuch finden können in dem ich ~~mir~~ meine geheimsten || geheimen Gedanken aufzeichne (in das ich sie sozusagen hineindeute).

Ms-115,88[3]et89[1] «verschiedene Rollen, die die Unwahrheit in Sprachspielen spielen kann» is replaced by «die Rollen unwahrer Sätze in Sprachspielen»

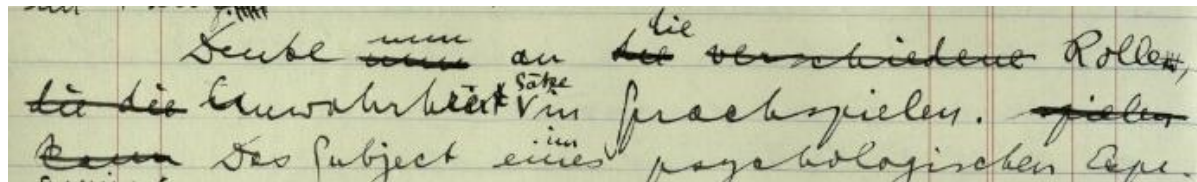


```
<choice type="em"><orig type="em1"><del type="d">die</del> <add rend="i">die</add>
  <del type="d">verschiedene</del> Rolle<del type="d"><del type="d_c">n</del></del>, <lb/> <del type="d">die
die</del>
<choice type="o"><orig type="o1">U</orig><orig type="o2">u</orig></choice>nwahrh<choice type="o"><orig
type="o1">ei</orig><orig type="o2">rer</orig></choice><del type="d">t</del> <add rend="im">Sätze</add> in
Sprachspielen<add rend="el">.</add> <del type="d">spielen<lb/> kann</del></orig>
```

```
<orig type="em2"><choice type="dsl"><orig type="alt1">die verschiedene Rolle, die die Unwahrheit in Sprachspielen
spielen kann<corr type="tra">.</corr></orig>
  <orig type="alt2">verschiedene Rollen, die die Unwahrheit in Sprachspielen spielen kann<corr
type="tra">.</corr></orig>
  <orig type="alt3">die Rolle, die die Unwahrheit in Sprachspielen spielen kann<corr type="tra">.</corr></orig>
  <orig type="alt4">die Rollen unwahr<corr type="npcn">h</corr><corr type="npcn">r</corr>er Sätze in
Sprachspielen.</orig></choice></orig></choice>
```



# Ms-115,88[3]et89[1]



Denke nun an die <sup>die</sup> verschiedene Rollen,  
die die Unwahrheit <sup>Sätze</sup> in Sprachspielen spielen.  
Das Subject eines <sup>im</sup> psychologischen Expe-  
rimentes soll sagen, was es gesehen hat; z.B. – es beschreibt seine Erfahrung falsch. – Der Meteorologe macht eine Prognose des zukünftigen Wetters; sie trifft nicht ein.

89

89

Der falsche Zug in diesem Sinne gehört wesentlich zum Spiel; er ist nicht eine Verunreinigung des Spiels, wie ein falscher Schritt im Tanz[.]en

nun

Denke nun an die <sup>die</sup> verschiedene Rollen<sup>e</sup>,  
die die [U|u]nwa<sup>r</sup>h[ei]r[er]t <sup>Sätze</sup> in Sprachspielen[.] spielen  
kann Das Subject eines <sup>im</sup> psychologischen Expe-  
rimentes <sup>Experiment</sup> soll sagen, was es gesehen hat; z.B.  
– es beschreibt seine Erfahrung falsch. – Der  
Meteorologe macht eine Prognose des zu-  
künftigen Wetters; sie trifft nicht ein.

Der falsche Zug in diesem Sinne gehört wesentlich zum Spiel; er ist nicht eine Verunreinigung des Spiels, wie ein falscher Schritt im Tanzen || Tanz.

Denke nun || nun an die verschiedene Rolle, die die Unwahrheit in Sprachspielen spielen kann. || verschiedene Rollen, die die Unwahrheit in Sprachspielen spielen kann. || die Rolle, die die Unwahrheit in Sprachspielen spielen kann. || die Rollen unwahrer Sätze in Sprachspielen. Das Subjekt eines psychologischen Experimentes || im psychologischen Experiment soll sagen, was es gesehen hat; z.B. – es beschreibt seine Erfahrung falsch. – Der Meteorologe macht eine Prognose des zukünftigen Wetters; sie trifft nicht ein.

89

Der falsche Zug in diesem Sinne gehört wesentlich zum Spiel; er ist nicht eine Verunreinigung des Spiels, wie ein falscher Schritt im Tanz.

Denke nun an die Rollen unwahrer Sätze in Sprachspielen. Das Subjekt im psychologischen Experiment soll sagen, was es gesehen hat; – es beschreibt seine Erfahrung falsch. – Der Meteorologe macht eine Prognose des zukünftigen Wetters; sie trifft nicht ein.

# Acknowledgment

The Wittgenstein Nachlass facsimile screenshots come from BNE 2015-:

- Wittgenstein Source Bergen Nachlass Edition. Edited by the Wittgenstein Archives at the University of Bergen under the direction of Alois Pichler. Bergen 2015-. In: Wittgenstein Source, curated by Alois Pichler (2009-) and Joseph Wang-Kathrein (2020-). <http://wittgensteinsource.org>
  - Facsimiles of MSS 105, 106 reproduced with the kind permission of © The Master and Fellows of Trinity College, Cambridge; The Österreichische Nationalbibliothek, Vienna; The University of Bergen, Bergen. All rights reserved.
  - Facsimiles of MSS 109, 114, 115, 122, 128 reproduced with the kind permission of © The Master and Fellows of Trinity College, Cambridge; The University of Bergen, Bergen. CC BY-NC 4.0.

# Further reading

- Claus Huitfeldt 2006: “Philosophy Case Study”. In: Electronic Textual Editing. Modern Language Association of America, 181-196. <https://tei-c.org/Vault/ETE/Preview/huitfeldt.html>
- Alois Pichler 2020: “A Typology of the Philosopher Ludwig Wittgenstein’s Writing of Text Alternatives”. In: Aisthesis: Pratiche, linguaggi e saperi dell’estetico, Vol. 13 / Number 2, 109-118. <https://oajournals.fupress.net/index.php/aisthesis/article/view/11166/11791>
- Alois Pichler 2021: “Hierarchical or Non-hierarchical? A Philosophical Approach to a Debate in Text Encoding”. In: DHQ: Digital Humanities Quarterly, Vol. 15 / Number 1. <https://digitalhumanities.org/dhq/vol/15/1/000525/000525.html>