

Kontatto/Kontatti Conditions of Use

1. Definitions

This document contains the **Conditions of Use and Code of Conduct (version 1.1 – April 10, 2018)** concerning the access, use, diffusion and publication of the data, services and other contents comprised in the following definitions, and it will be referred to as the **“Kontatto/Kontatti Conditions of Use”**.

1.1. “Kontatto”

The corpus **“Kontatto”** is a dataset collected by Silvia Dal Negro at the Free University of Bozen/Bolzano and/or under her supervision, during the research project “Italiano-tedesco: aree storiche di contatto in Sudtirolo e Trentino” (06.2011-05.2014) (hereafter, “Kontatto”). **Kontatto** consists of annotation files (mostly ELAN annotation files), audio files, metadata, questionnaires, map tasks and all the specific procedures and research protocols, including POS tagging, language and lemmatized form annotations. It contains the following subsets:

- **Kontatto-MT**: a subset with only the Map Task recording sessions of **Kontatto**
- **Kontatto-BA**: a subset of the recording sessions of **Kontatto** involving speakers from the “Bassa Atesina” area

Authorship and ownership of **Kontatto** is reserved to **Silvia Dal Negro** (Free University of Bozen/Bolzano) as coordinator of the above mentioned research project.

1.2. “Kontatti”

The corpus **“Kontatti”** is a dataset collected by the partners & collaborators of the research project “Germanico-Romanzo: discorsi e strutture in contatto nell’area dolomitica”, coordinated by Silvia Dal Negro (Free University of Bozen/Bolzano) and involving partners of the Università degli Studi di Trento (hereafter, “Kontatti”).

Kontatti consists of annotation files (mostly ELAN annotation files), audio files, metadata, questionnaires, map tasks and all the specific procedures and research protocols, including language and information structure annotations. It contains the following subsets:

- **Kontatti-A:** a subset of the corpus involving only speakers from the “Bassa Atesina” area
- **Kontatti-C:** a subset of the corpus involving only speakers of Cimbrian
- **Kontatti-L:** a subset of the corpus involving only speakers of (a variety of) Ladin
- **Kontatti-T:** a subset of the corpus involving only speakers of italo-romance varieties of Trentino

Authorship and ownership of **Kontatti** is reserved to **Silvia Dal Negro** (Free University of Bozen/Bolzano) as coordinator of the above mentioned research project.

2. *Kontatto*: Conditions of Use

The person who is being given access to *Kontatto* (hereafter, “**the User**”) declares that he/she is in agreement with the following conditions of use.

2.1. Access: general restrictions

Access and use of *Kontatto* is granted exclusively and individually to the User and in full compliance of the *Conditions of Use* detailed below, may be limited in time or otherwise restricted, and will be immediately terminated in case of non-compliance of these conditions.

2.2. Acknowledgement

The User shall always include a **clear acknowledgement of the source** in any research results, report or publication using data and/or contents originating from *Kontatto*.

The acknowledgement of the source shall always include:

1. **the name of the author/owner of the corpus:** “Silvia Dal Negro (Free University of Bozen/Bolzano)”
2. **the name and period of the research project:** “Italiano-tedesco: aree storiche di contatto in Sudtirolo e Trentino (06.2011-05.2014)”
3. **a reference to at least one of the publications listed in the related “References” section (2.9.)**
4. in substitution or addition to (3), **a reference to the internal project document** with the description of the corpus: “Dal Negro, Silvia / Ciccolone, Simone (2017), *Il corpus Kontatto: metodologia e descrizione*. Internal project document (URL: kontatti.projects.unibz.it/Kontatto-description)”

5. if applicable (e.g. for the subsets of the corpus submitted to *The Language Archive*), **the permanent digital identifier** of the dataset (see *The Language Archive* to get the persistent URL).

The User might use one of the following **examples** (as a footnote or within the text and adapting the text, language and reference format as required):

- The corpus *Kontatto* is a dataset owned by Silvia Dal Negro at the Free University of Bozen/Bolzano, collected during the research project “Italiano-tedesco: aree storiche di contatto in Sudtirolo e Trentino” (06.2011-05.2014). For further informations, see Dal Negro, Silvia / Ciccolone, Simone (2017), *Il corpus Kontatto: metodologia e descrizione*. Internal project document (URL: kontatti.projects.unibz.it/Kontatto-description); Dal Negro, Silvia / Ciccolone, Simone (2018), “Il parlato bilingue: italiano e tedesco a contatto in un corpus sudtirolese”. In: Bermejo Calleja, Felisa /Katelhön, Peggy (a cura di), *Lingua parlata. Un confronto fra l'italiano e alcune lingue europee*. Lang, Bern: 385-407.
- The data used in this study are part of the corpus *Kontatto*, collected during the research project “Italiano-tedesco: aree storiche di contatto in Sudtirolo e Trentino” (06.2011-05.2014) by Silvia Dal Negro (Free University of Bozen/Bolzano); cf. Dal Negro, Silvia / Ciccolone, Simone (2018), “Il parlato bilingue: italiano e tedesco a contatto in un corpus sudtirolese”. In: Bermejo Calleja, Felisa /Katelhön, Peggy (a cura di), *Lingua parlata. Un confronto fra l'italiano e alcune lingue europee*. Lang, Bern: 385-407.
- Corpus *Kontatto*, by Silvia Dal Negro (Free University of Bozen/Bolzano); research project: “Italiano-tedesco: aree storiche di contatto in Sudtirolo e Trentino” (06.2011-05.2014); cf. Dal Negro, Silvia / Ciccolone, Simone (2018), “Il parlato bilingue: italiano e tedesco a contatto in un corpus sudtirolese”. In: Bermejo Calleja, Felisa /Katelhön, Peggy (a cura di), *Lingua parlata. Un confronto fra l'italiano e alcune lingue europee*. Lang, Bern: 385-407.
- Corpus *Kontatto*, by Silvia Dal Negro (Free University of Bozen/Bolzano); research project: “Italiano-tedesco: aree storiche di contatto in Sudtirolo e Trentino” (06.2011-05.2014); cf. Dal Negro, Silvia / Ciccolone, Simone (2017), *Il corpus Kontatto: metodologia e descrizione*. Internal project document (URL: kontatti.projects.unibz.it/Kontatto-description).

2.3. Intended use

Kontatto may not be sold or used for commercial purposes. The access to the corpus is granted to the User only **for the sole purpose of conducting scientific research**.

In accordance to this condition, **(a) the User shall inform** the owner of *Kontatto* of all the collaborators with whom he/she intends to share or has shared the data and contents of *Kontatto*, which are thus tacitly deemed to agree with the same *Conditions of Use*; **(b) the User is invited** to notify any report, publication or presentation in which he/she and his/her collaborators intend to

use the data and contents included in *Kontatto*. In exchange, the owner of *Kontatto* agrees to share this information and references to the content in the project website.

2.4. Distribution of data

Kontatto and any part of its contents will be used only by the User, who is granted limited access to the corpus after returning an undersigned copy of the **Conditions of Use Agreement** (i.e. this document or any simplified version directly referring to this document).

Kontatto and any part of its contents may not, regardless of the form and/or medium, be further distributed, duplicated, circulated, made public or otherwise transmitted to any party without the prior consent of the author/owner of the corpus.

The User agrees to prevent unauthorized access and use of the data and contents in *Kontatto* by adopting appropriate security measures, such as protecting the credentials he/she uses to access the corpus, avoiding duplication of any part of its content if not necessary, in particular on shared devices or media, and exclusively on private access devices or media (protecting it with a secure password).

2.5. Personal data protection

The User undertakes to maintain confidentiality of all personal data that he/she processes and agrees to comply to general and specific privacy policy and conditions (such as anonymisation protocols, citation rules, etc.).

No information whatsoever (i.e., names, dates of birth, excerpts of the transcriptions and/or media files) enabling identification of informants shall be used and/or included in any report or communication.

When citing the materials (in the sole interest of illustrating his/her research), the User will adopt the speaker ID codes used throughout the dataset and remove or substitute identifying informations from the transcriptions and the audio clips used for the citation.

Materials contained in the corpus will not serve as the basis for personal judgments about the opinions, personality or language of the informants.

The User will not attempt to contact the informants, nor to interfere in any way in their personal lives.

Persons who use datasets containing personal data are required to comply with the Italian Legislative Decree 196/2003 covering Personal Data Protection ("**Codice in materia di protezione dei dati personali**", **D.Lgs. 196, 30/6/2003**) and its following changes and supplements.

2.6. Contribution

The User agrees to contribute to the corpus *Kontatto* in one or more of the following ways:

1. **through integration of the ELAN annotation files**, adding or completing annotation tiers concerning data analysis and information levels that might be of use for further researches (this does not necessarily include User-specific analysis annotations, but rather aims at enriching the corpus through largely re-usable annotations concerning segmentation, analysis and categorization of linguistic units, such as POS tagging, morphological tagging, lemmatization, syntactic tagging, information structure segmentation and tagging, phonetic and intonational tagging, etc.)
2. **by providing copies of the ELAN annotation files used for his/her research**, integrated with consistent additional tiers and annotations concerning data analysis and information levels that might be of use for further researches (as defined in the previous point)
3. **by donating new recording sessions** (each composed of one or more audio/video files of the recording session, one or more ELAN annotation files with basic transcription and at least the essential metadata as detailed in the *Kontatto* specifics) concerning one or more of the areas and languages of interest of *Kontatto* or *Kontatti* (areas: South Tyrol, Trentino, Ladin valleys; languages: South Bavarian dialects, Cimbrian, Ladin, Trentino dialects)

The User is also invited to contribute through revision and/or notification of errors and missing data in the transcription files.

All new content contributed by the User will be made available to other researchers in accordance with these *Conditions of Use*.

All contributions will be adequately acknowledged in the public domain descriptions and presentations of the corpus (i.e. the project website, the credit page on *The Language Archive*, future descriptions of the corpus, etc.).

2.7. Derivative works

Any derivative work of *Kontatto* (e.g. publications, conference talks, new datasets, data comparisons, reports etc.) is highly appreciated, provided that it fully complies to these Conditions of Use.

The User is therefore invited **(a) to inform the depositing party of his/her intended use of *Kontatto*** (course papers, thesis, conference talk, journal submission or other publication, web content, etc.), and **(b) to possibly provide some digital content concerning his/her work in**

whichever way he/she sees fit (e.g. abstract, pre-print version or final public domain publication, as well as notes, slide presentations, web link, etc.).

The content shared by the User may be added (with the appropriate acknowledgment) at his/her indication to an online repository concerning the project.

2.8. Changes in the Conditions of Use

The owner/author of *Kontatto* reserves the right, at his/her sole discretion, to modify or replace these Conditions of Use at any time. By continuing to access or use *Kontatto* after those revisions become effective, the User agrees to be bound by the revised conditions.

2.9. References for *Kontatto*

In accordance to Section 2.2, point 3, the User shall acknowledge the source of *Kontatto* by referring to at least one of the following publications:

- Dal Negro, Silvia (2013), “Dealing with bilingual corpora: parts of speech distribution and bilingual patterns”. *Revue française de linguistique appliquée* XVIII/2: 15-28.
- Dal Negro, Silvia / Ciccolone, Simone (2018), “Il parlato bilingue: italiano e tedesco a contatto in un corpus sudtirolese”. In: Bermejo Calleja, Felisa /Katelhön, Peggy (a cura di), *Lingua parlata. Un confronto fra l'italiano e alcune lingue europee*. Lang, Bern: 385-407.

3. Kontatti Conditions of Use

The person who is being given access to *Kontatti* (hereafter, “**the User**”) declares that he/she is in agreement with the following conditions of use.

3.1. Access: general restrictions

Access and use of *Kontatti* is granted exclusively and individually to the User and in full compliance of the *Conditions of Use* detailed below, may be limited in time or otherwise restricted, and will be immediately terminated in case of non-compliance of these conditions.

3.2. Acknowledgement

The User shall always include a **clear acknowledgement of the source** in any research results, report or publication using data and/or contents originating from *Kontatti*.

The acknowledgement of the source shall always include:

1. **the name of the author/owner of the corpus:** “Silvia Dal Negro (Free University of Bozen/Bolzano)”
2. **the name and period of the research project:** “Germanico-Romanzo: discorsi e strutture in contatto nell’area dolomitica (09.2016-08.2019)”
3. **a reference to the internal project document** with the description of the corpus: “Dal Negro, Silvia / Ghilardi, Marta (2018), *Il corpus Kontatti: metodologia e descrizione*. Internal project document (URL: kontatti.projects.unibz.it/Kontatti-description)”

The User might use one of the following **examples** (as a footnote or within the text and adapting the text, language and reference format as required):

- The corpus *Kontatti* is a dataset owned by the partners & collaborators of the research project “Germanico-Romanzo: discorsi e strutture in contatto nell’area dolomitica” (09.2016-08.2019), coordinated by Silvia Dal Negro (Free University of Bozen/Bolzano). For further informations, see Dal Negro, Silvia / Ghilardi, Marta (2018), *Il corpus Kontatti: metodologia e descrizione*. Internal project document (URL: kontatti.projects.unibz.it/Kontatti-description).
- The data used in this study are part of the corpus *Kontatti*, collected during the research project “Germanico-Romanzo: discorsi e strutture in contatto nell’area dolomitica” (09.2016-08.2019) coordinated by Silvia Dal Negro (Free University of Bozen/Bolzano); cf. Dal Negro, Silvia / Ghilardi, Marta (2018), *Il corpus Kontatti: metodologia e descrizione*. Internal project document (URL: kontatti.projects.unibz.it/Kontatti-description).
- Corpus *Kontatti*, by Silvia Dal Negro (Free University of Bozen/Bolzano); research project: “Germanico-Romanzo: discorsi e strutture in contatto nell’area dolomitica” (09.2016-08.2019); cf. Dal Negro, Silvia / Ghilardi, Marta (2018), *Il corpus Kontatti: metodologia e descrizione*. Internal project document (URL: kontatti.projects.unibz.it/Kontatti-description).

3.3. Intended use

Kontatti may not be sold or used for commercial purposes. The access to the corpus is granted to the User only **for the sole purpose of conducting scientific research**.

In accordance to this condition, **(a) the User shall inform** the owner of *Kontatti* of all the collaborators with whom he/she intends to share or has shared the data and contents of *Kontatti*, which are thus tacitly deemed to agree with the same *Conditions of Use*; **(b) the User is invited** to notify any report, publication or presentation in which he/she and his/her collaborators intend to use the data and contents included in *Kontatti*. In exchange, the owner of *Kontatti* agrees to share this information and references to the content in the project website.

3.4. Distribution of data

Kontatti and any part of its contents will be used only by the User, who is granted limited access to the corpus after returning an undersigned copy of the **Conditions of Use Agreement** (i.e. this document or any simplified version directly referring to this document).

Kontatti and any part of its contents may not, regardless of the form and/or medium, be further distributed, duplicated, circulated, made public or otherwise transmitted to any party without the prior consent of the author/owner of the corpus.

The User agrees to prevent unauthorized access and use of the data and contents in *Kontatti* by adopting appropriate security measures, such as protecting the credentials he/she uses to access the corpus, avoiding duplication of any part of its content if not necessary, in particular on shared devices or media, and exclusively on private access devices or media (protecting it with a secure password).

3.5. Personal data protection

The User undertakes to maintain confidentiality of all personal data that he/she processes and agrees to comply to general and specific privacy policy and conditions (such as anonymisation protocols, citation rules, etc.).

No information whatsoever (i.e., names, dates of birth, excerpts of the transcriptions and/or media files) enabling identification of informants shall be used and/or included in any report or communication.

When citing the materials (in the sole interest of illustrating his/her research), the User will adopt the speaker ID codes used throughout the dataset and remove or substitute identifying informations from the transcriptions and the audio clips used for the citation.

Materials contained in the corpus will not serve as the basis for personal judgments about the opinions, personality or language of the informants.

The User will not attempt to contact the informants, nor to interfere in any way in their personal lives.

Persons who use datasets containing personal data are required to comply with the Italian Legislative Decree 196/2003 covering Personal Data Protection ("**Codice in materia di protezione dei dati personali**", **D.Lgs. 196, 30/6/2003**) and its following changes and supplements.

3.6. Contribution

The User agrees to contribute to the corpus *Kontatti* in one or more of the following ways:

1. **through integration of the ELAN annotation files**, adding or completing annotation tiers concerning data analysis and information levels that might be of use for further researches (this does not necessarily include User-specific analysis annotations, but rather aims at enriching the corpus through largely re-usable annotations concerning segmentation, analysis and categorization of linguistic units, such as POS tagging, morphological tagging, lemmatization, syntactic tagging, information structure segmentation and tagging, phonetic and intonational tagging, etc.)
2. **by providing copies of the ELAN annotation files used for his/her research**, integrated with consistent additional tiers and annotations concerning data analysis and information levels that might be of use for further researches (as defined in the previous point)
3. **by donating new recording sessions** (each composed of one or more audio/video files of the recording session, one or more ELAN annotation files with basic transcription and at least the essential metadata as detailed in the *Kontatti* specifics) concerning one or more of the areas and languages of interest of *Kontatti* (areas: South Tyrol, Trentino, Ladin valleys; languages: South Bavarian dialects, Cimbrian, Ladin, Trentino dialects)

The User is also invited to contribute through revision and/or notification of errors and missing data in the transcription files.

All new content contributed by the User will be made available to other researchers in accordance with these *Conditions of Use*.

All contributions will be adequately acknowledged in the public domain descriptions and presentations of the corpus (i.e. the project website, the credit page on *The Language Archive*, future descriptions of the corpus, etc.).

3.7. Derivative works

Any derivative work of *Kontatti* (e.g. publications, conference talks, new datasets, data comparisons, reports etc.) is highly appreciated, provided that it fully complies to these Conditions of Use.

The User is therefore invited **(a) to inform the depositing party of his/her intended use of *Kontatti*** (course papers, thesis, conference talk, journal submission or other publication, web content, etc.), and **(b) to possibly provide some digital content concerning his/her work in whichever way he/she sees fit** (e.g. abstract, pre-print version or final public domain publication, as well as notes, slide presentations, web link, etc.).

The content shared by the User may be added (with the appropriate acknowledgment) at his/her indication to an online repository concerning the project.

3.8. Changes in the Conditions of Use

The owner/author of *Kontatti* reserves the right, at his/her sole discretion, to modify or replace these Conditions of Use at any time. By continuing to access or use *Kontatti* after those revisions become effective, the User agrees to be bound by the revised conditions.

4. Version history

| version | date | notes |
|---------|------------|--|
| 0.9 | 30.06.2017 | First full draft of the document |
| 1.0 | 10.07.2017 | First published version |
| 1.1 | 10.04.2018 | Minor revisions, update of the references for Kontatto |