

To be printed, filled, signed, scanned, and emailed to:

Dr. Keelan Evanini
660 Rosedale Road
Educational Testing Service
Princeton, NJ 08541

kevanini@ets.org

Please fill in the details in the sections highlighted in yellow in the document.

DATA LICENSE AGREEMENT

THIS AGREEMENT is effective as of the last signature date hereto (the “Effective Date”) by and between EDUCATIONAL TESTING SERVICE, a nonprofit, nonstock corporation organized and existing under the Education Law of the State of New York, with principal offices at Princeton, New Jersey 08541 (hereinafter “ETS”), and [REDACTED], a [researcher/professor/student] at [REDACTED] with the following address: [insert address] (hereinafter the “Researcher”).

WHEREAS, ETS is the owner of all right, title and interest and into certain data as described in the attached Exhibit A (hereinafter, the “Data”);

WHEREAS, Researcher desires access the Data for non-commercial research purposes, and ETS is willing to agree to such use of the Data, subject to the terms and conditions of this Agreement;

WHEREAS, subject to Researcher’s signature below, ETS’s acceptance of the terms and conditions set forth herein is deemed to have been provided upon ETS’s delivery of the Data to Researcher.

NOW, THEREFORE, in consideration of the foregoing premises and mutual covenants contained herein, ETS and Researcher, intending to be legally bound hereby, agree as follows:

1. LICENSE

- a) **Grant.** Subject to the terms and conditions of this Agreement, ETS hereby grants to Researcher a limited, non-exclusive, revocable, non-transferable license to use and reproduce the Data, and to publish results derived by Researcher from his or her analysis of the Data (“Results”), solely for the purposes of conducting non-commercial, academic research (“Non-Commercial Research”). Any other use or reproduction of the Data is prohibited.
- b) **Research Teams.** Researcher, in furtherance of the Non-Commercial Research, shall be permitted to make the Data available to other individuals, provided that (i) such individuals are acting under the direction and control of Researcher; (ii) such individuals may only use the Data in furtherance of the Non-Commercial Research being performed by Researcher; and (iii) Researcher shall be solely responsible for ensuring such individuals comply with the requirements of this Agreement including, without limitation, Sections 3, 4, and 6.
- c) **Attributions.** In the event that Researcher produces a written work that incorporates the Results or limited excerpts of the Data (with prior written consent from ETS), Researcher agrees to acknowledge the Data as property of ETS by displaying the following attributions:

Source: Derived from data provided by ETS

Copyright © 2016 ETS. www.ets.org

- d) **Disclaimer.** Researcher must also include the following disclaimer language in all publications:

“The opinions set forth in this publication are those of the author(s) and not ETS.”

2. FORMAT OF DATA

The Data will be provided to Researcher by ETS in the format described in Exhibit A.

3. **RESTRICTIONS.** The rights to the Data granted to Researcher pursuant to this Agreement shall be subject to the following restrictions:

- a) The license set forth in Section 1 shall permit research to be conducted only by, or under the direction and control of, the Researcher solely for non-commercial, academic research purposes. Access to the Data shall not be permitted to anyone except Researcher, except as expressly set forth in Section 1(b).
- b) Researcher may not reproduce the Data as a whole in any format. Researcher may reproduce limited excerpts of the Data solely in furtherance of Researcher’s non-commercial research.
- c) Except as otherwise expressly permitted by this Agreement, Researcher shall not publish, retransmit, display, redistribute, reproduce, or otherwise transfer the Data in any form to any third party.
- d) Researcher shall not commercialize, attempt to commercialize, or otherwise obtain commercial gain or benefit from his or her use of the Data.
- e) Researcher acknowledges and agrees that the Data he or she is receiving is not, and does not include, personally-identifiable data. Researcher agrees that he or she will not attempt to compile or match up Data with other personal data so that any individual(s) may be identified.
- f) Researcher shall use the Data solely for the purpose of and in accordance with this Agreement, and shall not otherwise use or disclose any of the Data to any third party.
- g) Researcher shall have no right to modify the Data and/or create any derivative works of the Data.

4. PROPRIETARY RIGHTS

- a) **Ownership of Data.** Except for the non-commercial, research license set forth in Section 1, ETS shall be and remain the sole and exclusive owner of all right, title, and interest in and to the Data, including all intellectual property rights therein, and no other rights or interests in the Data are transferred by or through this Agreement from ETS to Researcher, or any other third party.
- b) **Security and Notification.** Researcher shall ensure that the Data is secure and kept confidential at all times, and shall take necessary steps to prevent any unauthorized disclosure or use of such Data. Researcher shall immediately, upon discovery of any disclosure not authorized hereunder, notify ETS and take reasonable steps to prevent any further disclosure or unauthorized use.

- c) **Destruction of Data.** Upon termination or expiration of the Agreement, Researcher shall permanently destroy the Data and any and all copies thereof, and shall send written certification to ETS that all Data have been permanently destroyed.

5. EXCLUSION OF WARRANTIES

The Data provided by ETS to Researcher under this Agreement is “as is” and without any warranty, of any kind. ETS DISCLAIMS, AND RESEARCHER HEREBY WAIVES, ALL EXPRESS, IMPLIED AND/OR STATUTORY WARRANTIES, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT AND FITNESS FOR A PARTICULAR PURPOSE.

6. LIMITATION OF LIABILITY

In no event shall ETS be liable to Researcher or any individual receiving access to the Data pursuant to Section 1(b), regardless of the form of action or theory of recovery, in association with this Agreement, for (a) any indirect, special, exemplary, consequential, incidental or punitive damages, even if ETS has been advised of the possibility of such damages; (b) lost profits, lost revenue, lost business expectancy, benefit of bargain damages, business interruption losses or loss of data; or (c) direct damages or any kind.

7. TERM

- a) **Term.** This Agreement shall be effective as of the Effective Date and shall continue in full force and effect until terminated in accordance with this Section 7.
- b) **Termination.** ETS or Researcher may terminate this Agreement, for any reason, by providing written notice to each of the parties hereof of its intent to terminate this Agreement. In the event such notice is provided, termination of this Agreement shall be effective on the 30th calendar day following receipt of such notice.

8. IRREPARABLE HARM

The provisions of Sections 1, 3, and 4 of this Agreement may relate to unique and valuable assets of ETS. Breach or threatened breach of these sections may result in irreparable harm to ETS and remedies at law for such breach or threatened breach will be inadequate. In the event of a breach or threatened breach of Sections 1, 3, or 4, ETS shall be entitled to seek specific performance and/or injunctive relief, without any requirement to post a bond or other security. Such remedy may not be deemed to be the exclusive remedy for any such breach of this Agreement, but may be in addition to all other remedies available at law or equity.

9. REPRESENTATIONS AND WARRANTIES

- a) Researcher represents and warrants that he or she is empowered under applicable laws to enter into and perform this Agreement and that he or she has caused this Agreement to be executed and delivered to ETS.
- b) Researcher further represents and warrants that he or she will comply with any and all laws, rules and regulations, throughout the world, which are applicable to the performance of his or her obligations under this Agreement, including all applicable laws, rules, and regulations relating to privacy and data protection.

10. GOVERNING LAW

This Agreement shall be governed by the federal laws of the United States and the state laws of the State of New Jersey. In the event of any litigation related to this Agreement, the parties agree and consent to the exclusive jurisdiction of the State and Federal Courts located in the State of New Jersey, USA.

11. ASSIGNMENT

Researcher may not assign this Agreement without the prior written permission of ETS. Any attempt to assign any rights, duties, or obligations that arise under this Agreement without such permission shall be void.

12. SURVIVAL BEYOND TERMINATION OR EXPIRATION

The provisions of Sections 3 (Restrictions), 4 (Proprietary Rights), 5 (Exclusion of Warranties), 6 (Limitation of Liability), 8 (Irreparable Harm), and 10 (Governing Law), shall survive the termination or expiration of this Agreement.

EXHIBIT A

FORMAT AND DESCRIPTION OF DATA

Description of Data:

The ETS Corpus of Non-Native Spoken English is a new corpus of non-native English speech. It consists of spoken response provided during a high-stakes global assessment of English language proficiency, the Test of English as a Foreign Language (TOEFL iBT®). It contains 5,132 spoken responses to the first item in the Speaking section of the TOEFL iBT assessment sampled evenly from speakers of 11 different native languages (Arabic, Chinese, French, German, Hindi, Italian, Japanese, Korean, Spanish, Telugu, and Turkish). The responses were drawn from 7 different prompts (i.e., topics). The release of the TOEFL11 corpus is intended to support a broad range of research studies in the fields of spoken language processing and corpus linguistics. It was compiled with the specific task of Native Language Identification (NLI) in mind, but it will likely be useful for additional tasks and studies in the educational domain.

For the ComParE 2016 Native Language shared task, the ETS Corpus of Non-Native Spoken English corpus will be split as follows: 64% (3300 responses) will be supplied as training data, 19% (965 responses) as development data and 17% (867 responses) will be later released as test data.

Data Format:

The data will be released in the following format:

The audio files containing the spoken responses are in WAV format, mono, with 16 kHz sampling rate. A separate CSV file contains the metadata for each spoken response, including speaker ID, native language and prompt ID.