

PubPsych Metadata & Queries

– Translation Guidelines –

May 31, 2017

Abstract

This document should help you with the translation task for the CLUBS project. It tries to cover all the cases and some peculiarities you might find in the source material. If at any point in your work you are unsure of how to handle a specific translation situation, please *ask Roland prior to deciding for a variant*, because we probably want to include your case in these guidelines.

1 Query Translation

A query in PubPsych is a string of words with optional Boolean directives (AND, OR, NOT and round brackets) and specific fields (TI, LA...). The queries include (1) well-formed sentences, (2) ungrammatical sequences of words, and (3) a combination of any of them with other fields:

Examples

- (E1) Equity, excellence, and school reform: Why is finding common ground so hard?
- (E2) cognitive behavioral therapy physical trauma children
- (E3) process-dissociation models DB="PSYINDEX" LA="English"
- (E4) Handling pain OR stress AND (LA=French OR LA=German) AND PY>= 2003

1.1 Specific Guidelines

1. Translate all free text (that is text without a field label) and the text after the AB (abstract), CM (method type), IT (additional descriptor), KP (key phrase), SW (all keywords) and TI (title) fields. Ignore all other fields.

Original query: **process-dissociation models** DB="PSYINDEX" LA="English"
German translation: **Prozessdissoziation Modelle** DB="PSYINDEX" LA="English"

2. Translate text before and after Boolean directives and keep the Boolean directives as they are.

Original query: **Handling pain** OR **stress** AND (LA=French OR LA=German) AND PY>= 2003
German translation: **Umgang Schmerz** OR **Stress** AND (LA=French OR LA=German) AND PY>= 2003

3. Translate technical and domain-specific terminology (cf. E2) using the available resources with the following priority:
 1. built-in thesaurus of the Virtaal software,
 2. your previous translation memory,
 3. your new own translation.

Do not introduce additional words in your translation unless the same concept needs a different number of word in the two languages.

4. When translating grammatical sentences (cf. E1) try to preserve the original meaning and try to work with the least possible number of editing operations (word insertions, deletions and swaps) while keeping fluency in the target language.

FAQ

Can I use a dictionary?

Yes, you can use any resources you find suitable.

What should I do if there is a typo in the source query?

Translate it as if the typo does not exist and note the ID of the query on the appropriate paper form.

What should I do if there are truncations in the source query? How do I deal with wildcards?

Translate the query normally but keep the truncated word untranslated. Note the ID of the query on the appropriate paper form.

Which fields and their values should I ignore and not translate?

DT (document type), MT (media type), AGE (age group), EV (evidence level), PLOC (origin of population), AU (person(s)) ISBN, PU (publisher), ISSN, PY (publication year), COU (country of origin), SEG (database segment), CS (author affiliation), JT (journal title), SH (subject classification), CT (controlled term), DB (data source), LA (publication language)

What should I do if there is text following a field I don't recognise?

Please contact Roland if you find a field not mentioned here.

Should I translate named entities (names of persons, organizations, locations)?

Yes, translate them if applicable. For example, the French and Spanish *Londres* should become *London* in English or German. Most of the named entities are not going to be translated. Be careful with organizations, non-English institutions might have an English translation, such as *Leibniz-Zentrum für Psychologische Information und Dokumentation* in German would be *Leibniz Institute for Psychology Information* in English.

How should numbers, money amounts and dates be handled?

Please keep numbers as they are in the source. Numerals, like 4 or VII, should not be translated. If the numbers are spelled out, like *four*, they should be translated accordingly. Expressions containing numbers with a language dependent format, like money amounts or dates, should be formatted to the conventions of the target language.

2 Title and Abstract Translation

The abstract is the summary of an article available in the PubPsych database and it is composed by its title and the synopsis of the paper.

Examples

(E1) Testing Theories of Recognition Memory by Predicting Performance Across Paradigms

Signal-detection theory (SDT) accounts of recognition judgments depend on the assumption that recognition decisions result from a single familiarity-based process. However, fits of a hybrid SDT model, called dual-process theory (DPT), have provided evidence for the existence of a second, recollection-based process. In 2 experiments, the authors tested predictions of DPT and SDT by comparing the invariance of parameter estimates between yes/no (Y/N) and 2-alternative forced-choice (2AFC) testing paradigms. Both experiments showed DPT recollection estimates in Y/N to be poorly correlated with recollection estimates in 2AFC. In Experiment 2, SDT predictions explained more variance than DPT predictions. The authors evaluate and discuss the extent to which each model possesses theoretical validity versus computational flexibility in curve fitting.

(E2) What Determines the Speed of Lexical Access: Homophone or Specific-Word Frequency? A Reply to Jescheniak et al. (2003)

A. Caramazza, A. Costa, M. Miozzo, and Y. Bi (2001) reported a series of experiments showing that naming latencies for homophones are determined by specific-word frequency (e.g., frequency of nun) and not homophone frequency (frequency of nun + none). J. D. Jescheniak, A. S. Meyer, and W. J. M. Levelt (2003) have challenged these studies on a variety of grounds. Here we argue that these criticisms are not well founded and try to clarify the theoretical issues that can be meaningfully addressed by considering the effects of frequency on homophone production. We conclude that the evidence from homophone production cannot be considered to provide support to 2-layer theories of the lexical system.

2.1 Specific Guidelines

1. In general, try to preserve the original meaning and to introduce the least possible number of editing operations (word insertions, deletions and swaps) while keeping fluency in the target language.
2. Whenever possible, use the available resources with the following priority:
 1. built-in thesauri,
 2. your previous translation memory,
 3. your new own translation.
3. Translate titles with the following capitalisation rules: capitalize all nouns, pronouns, verbs, adjectives, and adverbs; lowercase articles, prepositions and coordinating conjunctions.

original query: **Behavioral Assessment and Interventions in Youth Sports**
 German query: **Verhaltensbewertung und Interventionen im Jugendsport**

4. Translate titles and abstracts sentence by sentence without any further splitting/join.

FAQ

Can I use a dictionary?

Yes, you can use any resources you find suitable.

What to do if there are typographic errors in the title or abstract?

There should be no errors. Please report any typographic error to Roland.

Should I translate bibliographic references?

No.

Should I translate acronyms?

Yes, acronyms should be translated if possible. For example, ADHD in English becomes ADHS in German. Another example is the acronym for human immunodeficiency virus: in English and German, it is HIV, whereas in French and Spanish it is VIH. Uncommon acronyms introduced by the authors of the abstract should not be translated, e.g. *essential amino acids (EAAs)* in English will be *essentielle Aminosäuren (EAAs)* in German.

Should I translate named entities (names of persons, organizations, locations)?

Yes, translate them if applicable. For example, the French and Spanish *Londres* should become *London* in English and German. Most of the named entities are not going to be translated. Be careful with organizations, non-English institutions might have an English translation, such as *Leibniz-Zentrum für Psychologische Information und Dokumentation* in German would be *Leibniz Institute for Psychology Information* in English.

How should numbers, money amounts and dates be handled?

Please keep numbers as they are in the source. Numerals, like 4 or VII, should not be translated. If the numbers are spelled out, like *four*, they should be translated accordingly. Expressions containing numbers with a language dependent format, like money amounts or dates, should be formatted to the conventions of the target language.

How should English compounds be handled?

Translate them as natural as possible without forcing a compound in the target language, e.g. *brain-intact animal* in English will be *Tier mit unversehrtem Gehirn* in German.