

多言語世界の扉を開く翻訳技術

人間の翻訳と機械の翻訳は何が違うのか？

影浦峽@「みんなの翻訳」運営チーム

<http://trans-aid.jp/>

NII市民講座 2010年6月3日

お話しの概要

- 翻訳はそもそもどんな行為か(プラトニックに)(1)
- 機械の「翻訳」メカニズムと現状
- 翻訳はそもそもどんな行為か(プラトニックに)(2)
- 機械翻訳の精度が仮にもっと向上したら？
- 結局は、コミュニケーションの原理に関わる問題
- 翻訳・情報流通・多言語バリアフリー：人間と機械の分業と協調

翻訳とはそもそもどんな行為か(プラトニックに)(1)

1. さっそくですが、次の文章を訳してみましょう:

나는 대학에 근무합니다.

2. 次はどうでしょうか?

Jo treballo per a una universitat.

3. もう一つ...

Mimi kazi kwa chuo kikuu cha.

4. それではこれはどうでしょうか?

I work at a university.

翻訳とはそもそもどんな行為か(プラトニックに)(1)

1. 訳せた人?

나는 대학에 근무합니다.

2. 訳せた人?

Jo treballo per a una universitat.

3. 訳せた人?

Mimi kazi kwa chuo kikuu cha.

4. 訳せた人?

I work at a university.

翻訳とはそもそもどんな行為か（プラトニックに）（1）

- では、今度はこれを訳してみましよう。

States Parties shall assure to everyone within their jurisdiction effective protection and remedies, through the competent national tribunals and other State institutions, against any acts of racial discrimination which violate his human rights and fundamental freedoms contrary to this Convention, as well as the right to seek from such tribunals just and adequate reparation or satisfaction for any damage suffered as a result of such discrimination.

翻訳とはそもそもどんな行為か(プラトニックに)(1)

- 皆さんが、今、訳せたものを訳した行為は、翻訳か？
- X文Y訳と翻訳の違いはどこにあるか？
 - X文Y訳は言語に関わる行為である
 - 翻訳は、いささかも言語に関わる行為ではない
 - 「バイリンガル」だから翻訳できるとは限らない
- では、言語とは何か？
 - 日本語、英語、スワヒリ語、カタロニア語・・・(言語学が扱う)
 - 言語表現／テキスト／メディアに記録され流通するもの
- これらの点には、あとで戻ってきます

お話しの概要

- 翻訳はそもそもどんな行為か(プラトニックに)(1)
- 機械の「翻訳」メカニズムと現状
- 翻訳はそもそもどんな行為か(プラトニックに)(2)
- 機械翻訳の精度が仮にもっと向上したら？
- 結局は、コミュニケーションの原理に関わる問題
- 翻訳・情報流通・多言語バリアフリー：人間と機械の分業と協調

機械の「翻訳」メカニズムと現状

1. 나는 대학에 근무합니다.

私は大学に勤務しています (G) / 私は大学に勤めます (E)

2. Jo treballo per a una universitat.

私は大学のために働く (G) / NA (E)

3. Mimi kazi kwa chuo kikuu cha.

私は大学のために働く (G) / NA (E)

4. I work at a university.

私は大学で働いています (G) / 私は大学で働いています (E)

機械の「翻訳」メカニズムと現状

締約国は、すべての人にその管轄効果的な保護および救済以内に国内裁判所およびその他の国の機関を通じて、彼の人権及び基本的自由を、この条約と同様、右追求するに反して侵害人種差別の行為に対する保証しなければならないただ、十分な補償や満足度など法廷損害そのような差別の結果として被ったから。(G)

機械の「翻訳」メカニズムと現状

州の党は彼らの管轄有効保護と療法の中で皆に保証されるものとします、有能な国家の裁判所と他の州団体を通して、この Conventionとは逆に彼の人権と基本的自由に違反する人種差別のどんな行為に対して、どんな損害に関する正当で適切な償いか満足もそのような区別の結果、受けたそのような裁判所から探す権利と同様に。(E)

機械の「翻訳」メカニズムと現状

- 以上から、わかること・・・
 - 300人より扱う言語の範囲は広い？
 - 簡単な文は訳せそう？
 - 難しくなると支離滅裂になる？
 - ?
 - ?
 - ?
 - ?

機械の「翻訳」メカニズムと現状：規則に基づく翻訳

- 二つの言語の辞書と文法規則・対訳辞書 (Excite翻訳)
- 例：
 - 「私はおじさんです。」← 辞書と文法規則で構文を解析
 - 私／は／おじさん／です。← 構造を言語に依存しない表現に
 - [私:主部][おじさん／だ:述部] ← 対訳辞書を適用
 - [I:主部][a ojisan:述部] ← 構造を英語のものに
 - “I am a ojisan.”

機械の「翻訳」メカニズムと現状：規則に基づく翻訳

- 困ること...

「形容詞は名詞にかかる」という日本語の規則

- 「太った山本さんの猫」

太っているのは山本さん、それとも、猫？

「前置詞句」は述語にかかる／名詞にかかるという二つの規則

- He saw a woman with a telescope.

望遠鏡で覗き見した？ 女性が望遠鏡を持っていた？

意味の曖昧性

- He went to the bank and got the money.

- He went to the bank and picked up morels.

機械の「翻訳」メカニズムと現状：用例に基づく翻訳

- 二つの言語のたくさんの（対訳）用例と辞書
- 例：
 - 「私はおじさんです。」←「私は銀行員です=I am a banker」例有り。
 - 私は[おじさん]です。←「おじさん=ojisan」辞書有り。
 - “I am an ojisan”

機械の「翻訳」メカニズムと現状：規則に基づく翻訳

- 困ること・・・
 - 言語データは稀なので用例が当たりにくい
 - 単語だけで数十万～数百万、単語の並びだと？

機械の「翻訳」メカニズムと現状：統計翻訳

- 二つの言語のたくさんの(対訳)用例(と辞書)からつながりと対応の確率を計算しておく(Google翻訳)
 - 一つの言語のつながりの確率：言語モデル
 - 対応の確率：翻訳モデル
- 困ること
 - 構造が異なる言語のときは特に、構文が無茶苦茶になることが少なくない

機械の「翻訳」メカニズムと現状

- 以上から、わかること・・・
 - 300人より扱う言語の範囲は広い
 - 簡単な文は訳せそう
 - 難しくなると支離滅裂になる
 - 機械翻訳は「言語」に関わるプロセスである
 - ?
 - ?
 - ?

機械の「翻訳」メカニズムと現状

- 中学生の方？
- 小学生の方？
- 研究者の方？
- 翻訳のプロの方？
- 日本語を母語としていない方？
- Google翻訳／Excite翻訳で同じ文を翻訳して比べてみてください。

機械の「翻訳」メカニズムと現状

- 以上から、わかること・・・
 - 300人より扱う言語の範囲は広い
 - 簡単な文は訳せそう
 - 難しくなると支離滅裂になる
 - 機械翻訳は「言語」に関わるプロセスである
 - 相手に応じて訳を変えることができない
 - ?
 - ?

お話しの概要

- 翻訳はそもそもどんな行為か(プラトニックに)(1)
- 機械の「翻訳」メカニズムと現状
- 翻訳はそもそもどんな行為か(プラトニックに)(2)
- 機械翻訳の精度が仮にもっと向上したら？
- 結局は、コミュニケーションの原理に関わる問題
- 翻訳・情報流通・多言語バリアフリー：人間と機械の分業と協調

翻訳とはそもそもどんな行為か(プラトニックに)(2)

- 質問: 本当にプロの「翻訳者」はこの文を前にどうするでしょうか?

States Parties shall assure to everyone within their jurisdiction effective protection and remedies, through the competent national tribunals and other State institutions, against any acts of racial discrimination which violate his human rights and fundamental freedoms contrary to this Convention, as well as the right to seek from such tribunals just and adequate reparation or satisfaction for any damage suffered as a result of such discrimination.

翻訳とはそもそもどんな行為か(プラトニックに)(2)

- 質問: 本当にプロの「翻訳者」はこの文を前にどうするでしょうか?
 1. この限りでは翻訳はできません、と言う。
 2. 一応、この文を含む文章の翻訳を受けていた場合は・・・
 - 既訳とりわけ「正訳」を探す(条約の一部であることがわからなければプロではない)
 - 対象読者がこの文書と関連づける言語表現域を確定
 - 「正訳」をそのまま使うか、対象に応じてあえて表現を変えるかする

翻訳とはそもそもどんな行為か(プラトニックに)(2)

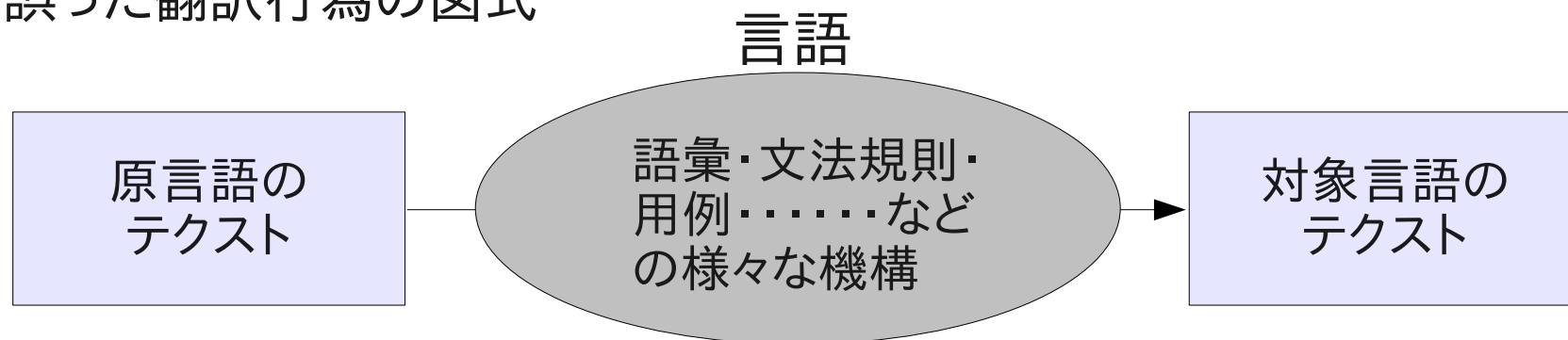
- 翻訳は「言語」に関わる行為ではなく「言語表現」に関わる行為
 - 誰かがどこかで発話しない限り「言語表現」は存在しない
 - 「言語表現」は歴史的・社会的な存在
 - 「言語」はそこから歴史・社会性を剥奪して形を抽出したもの
 - 私たちは「日本語」を話しているといったとき、誰も「日本語」なるものを見たことも聞いたこともない
 - 見たり聞いたりするのは個々の「表現」
 - 翻訳は「言語表現」を歴史・社会的な相のもとで扱う

翻訳とはそもそもどんな行為か（プラトニックに）（2）

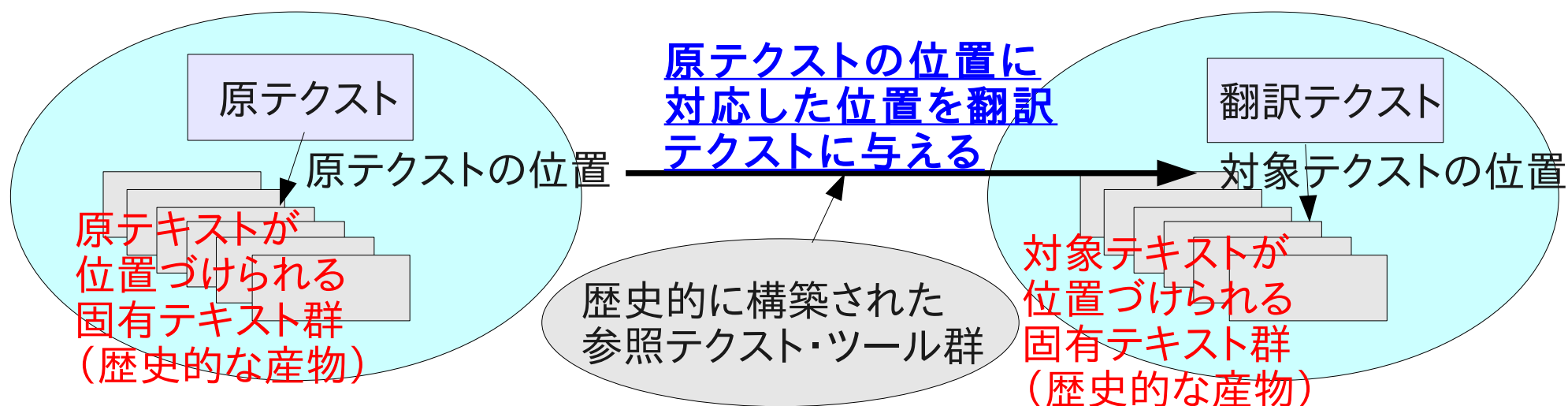
- 翻訳は、あえて関連する他の行為と差異化する点から規定するならば、いかなる意味でも、いささかも認知的あるいは言語的行為ではなく、テキスト／言語表現を、社会に流通可能かつ蓄積参照可能なかたちで提出する、徹頭徹尾社会的な行為である。
 - 言語学的な意味での言語も意味も「翻訳」がそれ以外の何でもなくまさに「翻訳」であることにはいささかも関わりはしない。
- 翻訳元のテキストは、その言語において蓄積されてきたテキストの織物の中に位置づけられるべく公表され社会的に蓄積記憶される。翻訳は、翻訳元のテキストが原言語におけるテキストの織物に占める位置に相当する対象言語におけるテキストの織物上の位置に翻訳テキストを位置づける行為である。

翻訳とはそもそもどんな行為か(プラトニックに)(2)

- 誤った翻訳行為の図式



- 概ね正しい翻訳行為の図式



翻訳とはそもそもどんな行為か(プラトニックに)(2)

言語の分析が、ある言説の事実に関して問う問いとはつねに、

「どのような規則にもとづいて、このような言表が作られたのか？ またしたがって、どのような規則によって他の同じような言表をつくることができるのか？」

であるのに対して、言説の記述が立てるのはまったく別の問いであって、

「このような言表が出現した、しかも、他のいかなる言表もその代わりには出現しなかったのは、どのようなわけなのか？」

という問いなのだ。

翻訳とはそもそもどんな行為か(プラトニックに)(2)

- 扱うもの
 - 翻訳: 言語表現／テキスト(歴史社会的な産物)
 - 機械翻訳: 言語として見た言語表現(歴史社会性を剥奪)
- 立場
 - 翻訳: 翻訳した文書の歴史社会的な位置を定める意思決定
 - 機械翻訳: 計算
- 翻訳の問題点
 - 「対応する位置」を定めると言っても、そもそも対応する目安がない場合はどうするのでしょうか？ これについては、後ほど

翻訳とはそもそもどんな行為か(プラトニックに)(2)

- フィクションでありうる極端な例(村上龍『69』)

…わかった、[浅丘ルリ子](#)よりきれいに映さんばだめばい、と言って、講堂の入口付近の最前列に並んでいた連中の尻を蹴っ飛ばし、場所を空けさせた。松井和子が主演女優ということで、裕二は急に上機嫌になり、主題歌は石原裕次郎がいいとか、…

... "Yeah, that makes sense. But listen, man, you better make her look good. Better than [Brigitte Bardot](#), even." He moved up to the front of the crowd, kicking people in the butt to clear a space for us. The idea of making Kazuko Matsui the star of our film had excited him, and now he was rattling on about how we should use something by Yujiro Ishihara for the theme song, ...

機械翻訳の精度が仮にもっと向上したら？

- X文Y訳としての精度向上
 - これが機械翻訳研究がやってきたこと
 - これはまだそこそこあがりそう……
- 「翻訳」として存在するために
 - あるところからは精度の問題ではない
 - その意味では、精度がどんどん向上してもダメ
 - 自己認識とその自己認識に対するユーザの理解
 - いつ、どこで、どんなときに、それなりに使えるか？

お話しの概要

- 翻訳はそもそもどんな行為か(プラトニックに)(1)
- 機械の「翻訳」メカニズムと現状
- 翻訳はそもそもどんな行為か(プラトニックに)(2)
- 機械翻訳の精度が仮にもっと向上したら？
- 結局は、コミュニケーションの原理に関わる問題
- 翻訳・情報流通・多言語バリアフリー：人間と機械の分業と協調

結局は、コミュニケーションの原理に関わる問題

- 提起されたけど検討されていない点：
 - 翻訳対象側に翻訳文書を位置づける既往文書群がないとき
 - 機械翻訳の「自己認識」
- よく言われる点：
 - 文化の違いは翻訳できない、などなど

結局は、コミュニケーションの原理に関わる問題

- 「(文化が違えば)翻訳は不可能です」と言うのは間違っている
 - 神学論争や勝手な思い込みではなくそれなりに検証し検討することのできる主張としては「翻訳は不可能です」と言えるために、可能性としては無限にありうる翻訳それぞれにつき「それは正しい翻訳ではありません」とそれなりに検証し検討することのできるかたちで言えなくてはならない。そのためには「正しい翻訳」の存在が論理的に要請される(少なくとも統整的な理念として)。
 - 誰一人として同じ経験をしている人はおらず、同じ言葉を話している人もいない。したがって「(文化が違えば)翻訳は不可能です」という主張は、結局、誰であれ人は自分以外とは決して分かり合えないという主張に行き着く(仮に同じ「言語」を話していても。少なくとも私自身は「(文化が違えば)翻訳は不可能です」と主張する日本人よりは「翻訳できるかも」というフランス人やケニア人との方がより分かり合える)。

結局は、コミュニケーションの原理に関わる問題

- 翻訳対象側に翻訳文書を位置づける既往文書群がないとき
 - 個別の対処方法(『シチリアでの会話』)
 - 全体としては……
 - ゆっくりと時間をかけて慣れていくこと
(cf. 赤ん坊が最初に言葉を習得するとき)
 - 影浦峽『子どもと話す 言葉って何?』(現代企画室, 2006年, 1200円+税)
- 機械翻訳の「自己認識」
 - これから「自己認識」を技術的に展開するイメージ
 - 現状に「自己認識」を加えるイメージで考えると?

お話しの概要

- 翻訳はそもそもどんな行為か(プラトニックに)(1)
- 機械の「翻訳」メカニズムと現状
- 翻訳はそもそもどんな行為か(プラトニックに)(2)
- 機械翻訳の精度が仮にもっと向上したら？
- 結局は、コミュニケーションの原理に関わる問題
- 翻訳・情報流通・多言語バリアフリー:人間と機械の分業と協調

翻訳・情報流通・多言語バリアフリー

- 機械翻訳は、一部の多言語バリアフリー化に、使えます
 - ホテルの情報を自分で確認する
 - ロシアのオンラインサイトから商品を購入する
 - 人間の翻訳者は、普段どんなに威張って「機械翻訳はダメだ」と言っている、これらの状況にいる人の手伝いは決してしてくれません。機械翻訳なら、文句を言わずやってくれます。
 - こうしたコミュニケーションの場とそこでのやりとりにも、必要があればゆっくり時間をかけて馴染んでいくことができます。
 - 「歓待とは忘却の振舞いではなくあくまで差異に対する意識化の振舞い」です。

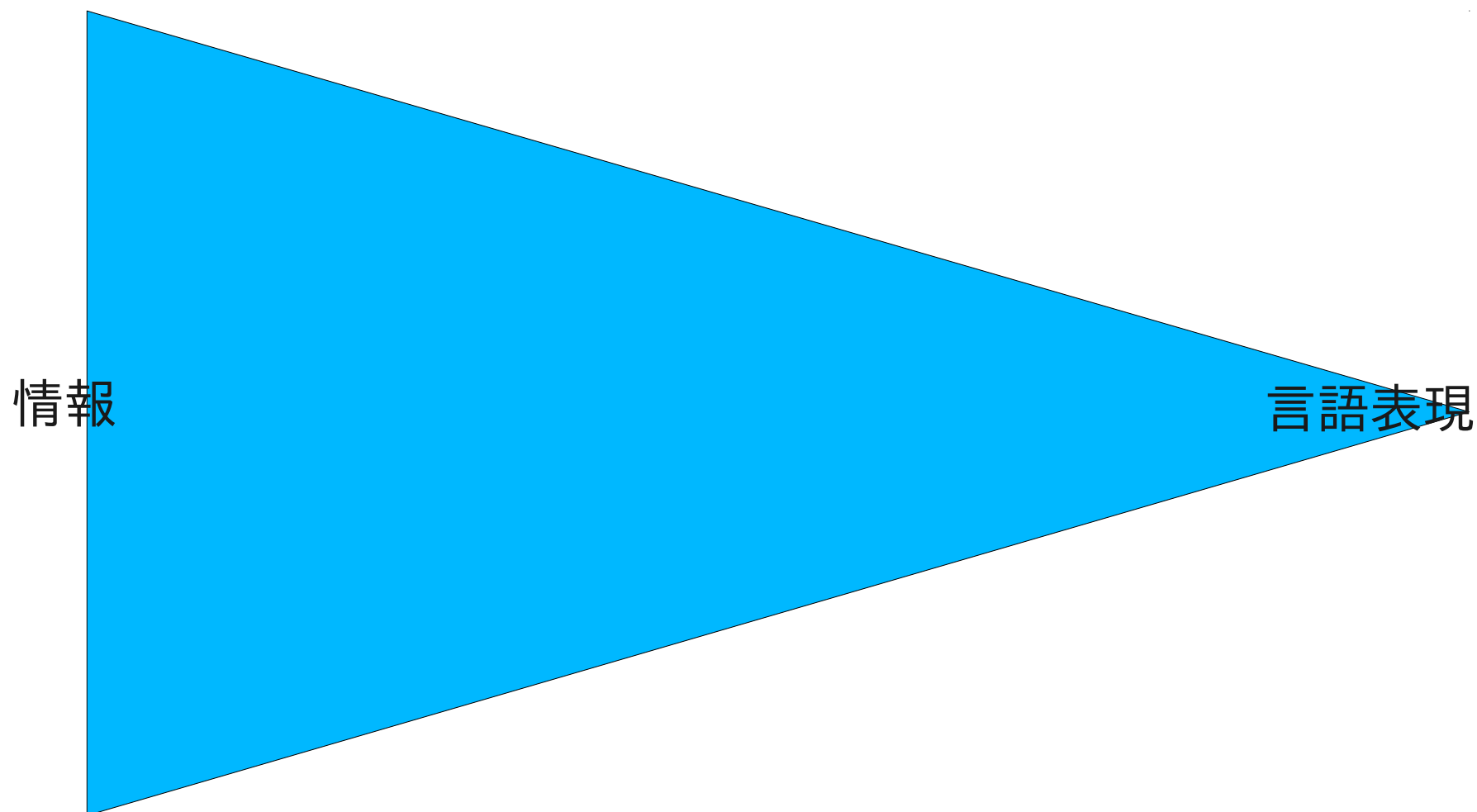
翻訳・情報流通・多言語バリアフリー

- 機械翻訳が精度の問題として使えないところもあります
 - 契約書の翻訳、条約の翻訳、などなど
 - その他広い範囲のちょっと複雑な、ほとんどのテキストで
 - 人間の翻訳者がやったものは使えます。ただし、トラブルが起きうるという意味では、常に正しいわけではありません。
 - 人間の翻訳を介したやりとりには、これまでゆっくり時間をかけて馴染んできたことから私たちの心構えも社会の制度もある程度できています。
 - 一方で、そのように「馴染んだ」人間の翻訳が扱える範囲はごくわずかです。

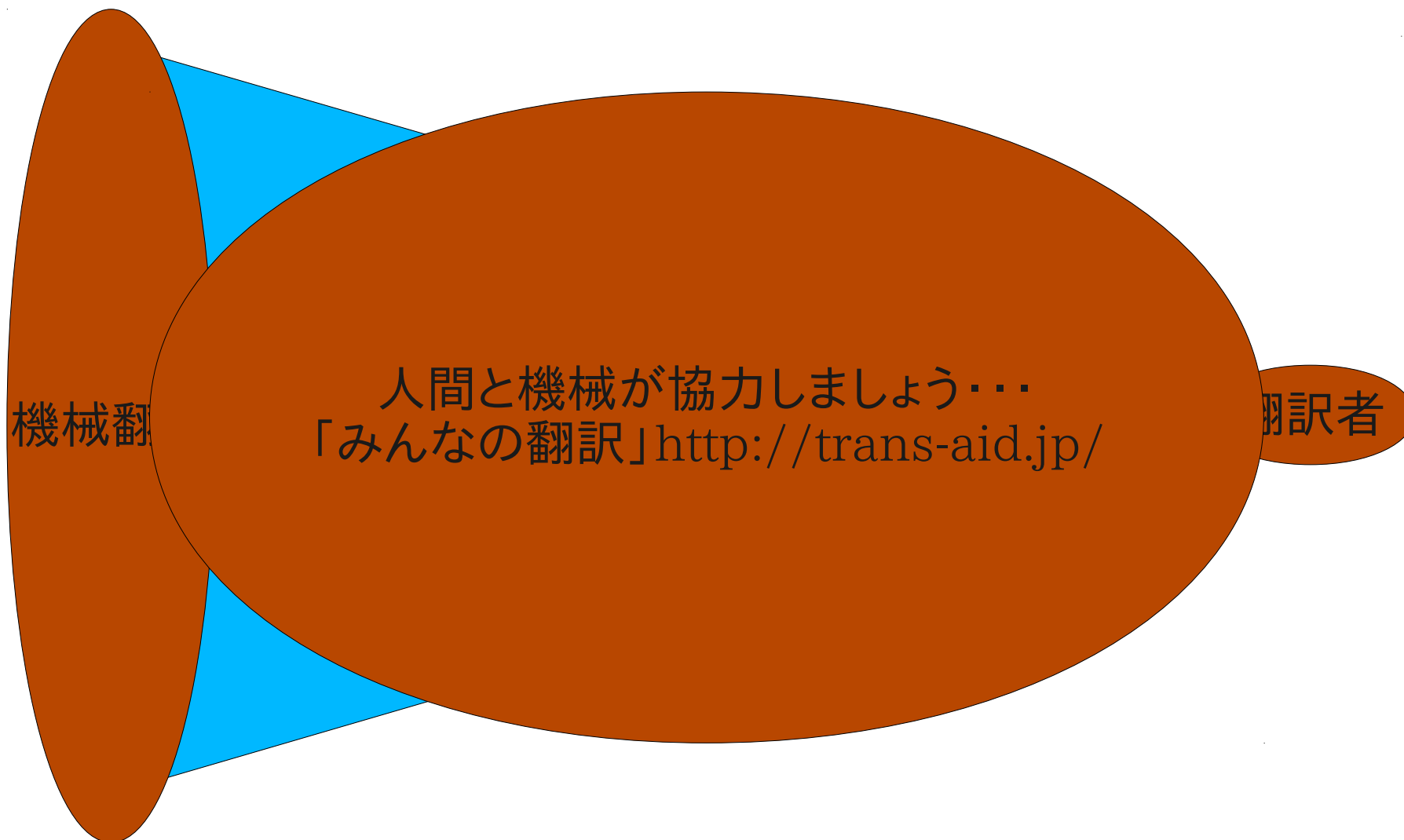
翻訳・情報流通・多言語バリアフリー

- 以前は人間以外が「翻訳」するなど思いもよらなかった
 - 多言語情報流通は限られた対象、限られた言語、限られた人の中で行われていた
 - メガネができる前は視力0.1の人は「バリア」に囲まれていた
 - 現在、技術は対象と言語と人の範囲を広げた／広げうるけれど、どこについて広げることができるか、皆、戸惑っている

翻訳・情報流通・多言語バリアフリー



翻訳・情報流通・多言語バリアフリー



扱わなかった問題

- 機械翻訳が今以上に精度を、自己認識とともに向上させる方向
- 人間と機械の協調と翻訳
- 人間を馬鹿にしたシステムと人間を賢くするシステム
- その他もろもろ