Language solutions for education in emergencies

Alice Castillejo: Program Advisor
Kate Murphy: Plain Language Advisor
Grace Tang: Program Manager, Gamayun
TWB is working towards a world where knowledge knows no language barriers

Language maps
Language technology
advocacy
Education glossaries
services

Translation
Research and
Plain language
<table>
<thead>
<tr>
<th>CONCEPTS (français)</th>
<th>Swahili</th>
<th>Kiswahili</th>
<th>Masi</th>
<th>Audre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maladie</td>
<td>OBUKONI</td>
<td>ndwalo</td>
<td>ndwalo</td>
<td></td>
</tr>
</tbody>
</table>
| Vaccin             | endug 
| Signes             | ECHIMIN 
| Consentement        | KALO | ERYAAN 
| Soins               | OBUBA | MWABUKA 
| Bonjour             | WASING 
| Merci               | Bwaken |
Plain language can change reader response

<table>
<thead>
<tr>
<th></th>
<th>Percentage of registrations in response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original invitation</td>
<td>43.3</td>
</tr>
<tr>
<td>Plain-language invitation</td>
<td>78.4</td>
</tr>
</tbody>
</table>
It’s not just about vocabulary

Substance

Plain language

Structure

Statements
Use the inverted pyramid model

- Essential point
- Strong evidence
- Extra information

Consistent peak message
Decreasing importance
Most readers scan text

Subheadings help readers to navigate

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam sit amet diam id eros sollicitudin sollicitudin. Aliquam ac vehicula purus. Cras tempus, est consequat elementum condimentum, elit felis aliquet est, ut interdum tellus orci fringilla justo.


Bullet points aid memory and recall

- Aenean enim ipsum, lacinia ac purus tincidunt, accumsan sodales est.
- Etiam aliquet aliquet mi at pharetra. Vivamus sed orci non diam condimentum hendrerit.
- Mauris vulputate, turpis fermentum gravida faucibus, justo erat porta est, in.

Polio pamphlet

![Bar chart showing comparison before and after editing](chart.png)

Figure 1: Plain language improves reading time

Graphics can make things clearer

Condimentum, elit felis aliquet est, ut interdum tellus orci fringilla justo.

It’s not just about vocabulary
Sentence length affects comprehension

Source: American Press Institute

TWB recommends an average sentence length of fewer than 20 words
Use familiar words

<table>
<thead>
<tr>
<th>Rank</th>
<th>Word/lemma</th>
<th>PoS</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>the</td>
<td>a</td>
<td>22038615</td>
</tr>
<tr>
<td>2</td>
<td>be</td>
<td>v</td>
<td>12545825</td>
</tr>
<tr>
<td>3</td>
<td>and</td>
<td>c</td>
<td>10741073</td>
</tr>
<tr>
<td>4</td>
<td>of</td>
<td>i</td>
<td>10343885</td>
</tr>
<tr>
<td>5</td>
<td>a</td>
<td>a</td>
<td>10144200</td>
</tr>
<tr>
<td>6</td>
<td>in</td>
<td>i</td>
<td>6996437</td>
</tr>
<tr>
<td>7</td>
<td>to</td>
<td>t</td>
<td>6332195</td>
</tr>
<tr>
<td>8</td>
<td>have</td>
<td>v</td>
<td>4303955</td>
</tr>
<tr>
<td>9</td>
<td>to</td>
<td>i</td>
<td>3856916</td>
</tr>
<tr>
<td>10</td>
<td>it</td>
<td>p</td>
<td>3872477</td>
</tr>
<tr>
<td>11</td>
<td>I</td>
<td>p</td>
<td>3978265</td>
</tr>
<tr>
<td>12</td>
<td>that</td>
<td>c</td>
<td>3430996</td>
</tr>
<tr>
<td>13</td>
<td>for</td>
<td>i</td>
<td>3281454</td>
</tr>
<tr>
<td>14</td>
<td>you</td>
<td>p</td>
<td>3081151</td>
</tr>
<tr>
<td>15</td>
<td>he</td>
<td>p</td>
<td>2909254</td>
</tr>
<tr>
<td>16</td>
<td>with</td>
<td>i</td>
<td>2683014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rank</th>
<th>Word/lemma</th>
<th>PoS</th>
<th>Part of Speech (PoS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>a</td>
<td>a</td>
<td>VERB</td>
</tr>
<tr>
<td>2201</td>
<td>abandon</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>783</td>
<td>ability</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>386</td>
<td>able</td>
<td>j</td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>abortion</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>about</td>
<td>i</td>
<td></td>
</tr>
<tr>
<td>179</td>
<td>about</td>
<td>r</td>
<td></td>
</tr>
<tr>
<td>896</td>
<td>above</td>
<td>i</td>
<td></td>
</tr>
<tr>
<td>1599</td>
<td>above</td>
<td>r</td>
<td></td>
</tr>
<tr>
<td>3392</td>
<td>abroad</td>
<td>r</td>
<td></td>
</tr>
<tr>
<td>2547</td>
<td>absence</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>3192</td>
<td>absolute</td>
<td>j</td>
<td></td>
</tr>
<tr>
<td>1152</td>
<td>absolutely</td>
<td>r</td>
<td></td>
</tr>
<tr>
<td>3145</td>
<td>absorb</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>4045</td>
<td>abstract</td>
<td>j</td>
<td></td>
</tr>
<tr>
<td>1548</td>
<td>abuse</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>come</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Multilingual emergencies

60 countries
4,984 languages
204,413,000 people in need
Language Technology

- Multi-lingual Speech Recognition
- Spoken Language Translation
- Multi-lingual Speech Synthesis

- Multi-speaker Speech Corpora
- Text Corpora of Source Language
- Parallel Corpora Text Corpora of Target Language
- Large Scale Speech Corpora of Target Language
Size of publicly available language datasets

- French: 3,800.0M
- Spanish: 2,300.0M
- German: 1,600.0M
- Russian: 1,000.0M
- Turkish: 900.8M
- Arabic: 802.3M
- Slovenian: 578.9M
- Persian: 125.8M
- Indonesian: 125.2M
- Vietnamese: 72.4M
- Urdu: 25.5M
- Bengali: 21.6M
- Kurdish: 7.9M
- Nepali: 5.4M
- Hausa: 3.1M
- Swahili: 2.7M
- Yoruba: 0.5M
- Igbo: 0.3M
- Kanuri: 0.0M
- Rohingya: 0.0M
Building Machine Translation

- Translate basic content (Gamayun core corpus)
- Record audio content
- Collect parallel data (TWB, other sources, etc)
- Develop and test model
- Develop interface
- Integration Value added
Humanitarian Action Challenge

[Image of a blue speech bubble with the WFP logo and "vam food security analysis" and a red brick wall separating two groups of people, one in blue and the other in grey. One blue person is speaking to the group with "Levantine Arabic speakers" speech bubble.]
Khabrona provides accurate and reliable information on the civil and legal information relevant to Syrian refugees living in Jordan.

- Supports user-base of >30% women
- Provides ~30,000 refugees with integration information
- Reaches users through Facebook and website
Opportunities

Effective 2 way communication

Interaction

Accountability

Increased capacity (speed)
Get Involved!

Partnerships
Data sharing
Volunteer
Donate
Use word frequency data

Plain language lexicon by Martin Cutts

Top 5,000 English words

www.testyourvocab.com
Email subject: Thank you for your application at Translators without Borders!

Dear NAME,

Welcome to Translators without Borders (TWB)!

You are now ready to get started as a Kató Translator. We have prepared this welcome pack to help you take your first steps with TWB.

**Your next steps:**
1. Go to Kató Platform and click on "Register";

2. Insert the email address you used on the sign-up form and choose your password;

3. Once you have activated your account, please sign in and complete your profile. You can start translating right away, if there are tasks available in your languages :)

If you need any support, please contact translators@translatorswithoutborders.org.

We look forward to working with you,

The TWB team
Dear <<First Name>>,

Welcome to Translators without Borders (TWB), and thank you for applying.

It’s easy to start translating with us. You just need to register on our online platform, Kató. This is where we post translation tasks.

**It will only take a minute:**

1. Open [Kató Platform](#) and click "★ Register"
2. Enter the email address you used on the application form, enter a password, then click “★ Register”
3. Open the email you receive from TWB Kató and click “Click here to finish registration”
4. Click “Finish Registration” on the page that opens
5. Sign in and complete your profile
6. Watch [this tutorial](#) to understand how to find available tasks in your languages
7. Start translating and using your language skills to help people in need!

We have prepared [this welcome pack](#) with more information on your first steps with TWB.

If you need any support, please contact translators@translatorswithoutborders.org.

Thanks for your interest in TWB! Your volunteer work will help people around the world access information in a language they can understand.

The TWB team