TransPHorm

**Trans**lation of **P**ublic **H**ealth Inf**orm**ation for a Diverse Public
University of Washington

# Addressing language disparities: Use of machine translation to improve access to multilingual health materials

Anne M. Turner, MD, MLIS, MPH
Associate Professor, University of Washington
Sydney, Australia
December 3, 2014

# Introduction

- Large number of individuals in US and Australia that speak a language other than English
- Individuals with limited English proficiency have less access to health care, preventative services and poorer health outcomes than English speaking minorities
- There is a tremendous needs for multilingual health materials
- This need is unmet due to the time and costs associated with producing quality translations

**Trans**lation of **P**ublic **H**ealth Inf**orm**ation for a Diverse Public
University of Washington

# Statistical Machine Translation

- The translation of text from one human language to another by a computer that is "trained" to translate based on vast amounts of previously translated text.

- Getting better all the time, but not perfect.

# Poor translations, but getting better

## 2010

- *Source:* Clean unwashable wallpaper with commercial putty-like wallpaper cleaner.
- *Translation:* Fondos de escritorio de Limpieza lavable con masilla un limpiador comercial fondo de pantalla.
- *Back translation:* Depths/backgrounds of desk of cleaning washable with putty a commercial cleaner screener background.

## 2014

- *Source:* Clean unwashable wallpaper with commercial putty-like wallpaper cleaner.
- *Translation:* Limpia fondos no lavables con papel pintado comercial de masilla-como limpiador.
- *Back translation:* Clean nonwashable wallpaper with PuTTY -like cleaner commercial  wallpaper.

# Linguistic error types:

**a. Missing term** (leave out a term essential to the meaning)

**b. Word order errors**

**c. Wrong word sense**

**d. Morphological errors** (wrong word form, plural/singular word mismatch) -    *"He putted the plate on the table."*

# Most annoying errors

| Error Category | % |
|---|---|
| **Morphological error** | **28.2** |
| Missing term | 16.7 |
| Word sense error | 16.1 |
| Word order error | 9.7 |
| Punctuation | 9.1 |
| Spelling | 5.1 |

# Methods I

- In-depth interviews to determine translation processes
- Task analysis, identify work processes, cost and time for manual translations

# Interviews

- In-depth interviews with personnel from five health departments (n=41)
- Identified barriers and facilitators
- Assessed tasks, goals, and attitudes towards machine translation

# Four Phases of Manual Translation

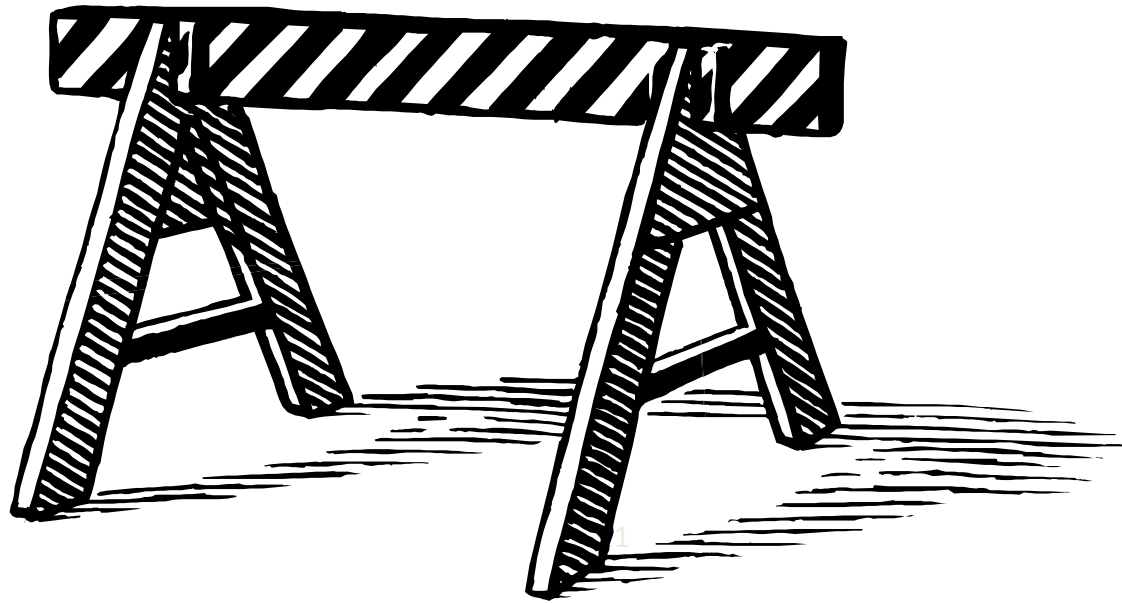Pre-Translation Phase

Translation Phase

Quality Assurance Phase

Post-Translation Phase

# Barriers to Manual Translation

o Cost

o Trained staff

o Time

o Lack of awareness

o Decentralization

# Time and Costs: Manual

| Manual Translations | | |
|---|---|---|
| Time |  | 17 hours-6 days |
| Cost |  | $130-$1,200 (US)<br>$152-$1,410 (Aust) |

# Attitudes toward machine translation

- Concerns about errors and quality
- Smaller health departments with fewer resources more favorable
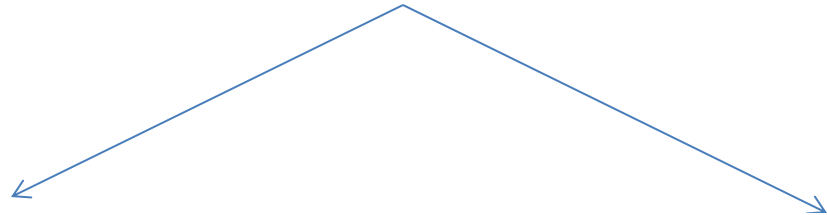- Concerns about job security

# Methods II

- Post-editing studies
- Comparison rating by experts

# Machine Translation Post-Editing Study

30 English health promotion documents from health departments



Manual translation by
human translator

Machine translation
+
post-editing by bilingual staff

# Rank Finished Product



Manual





MT +  Post - editing

# Rank Finished Product (n=30)

Manual
**10**

Equal
**10**

MT + Post – editing
**10**

# Summary Results: Spanish

| | Manual | MT + Post-edit |
|---|---|---|
| Time | 17 hrs -6 days | 3 min – 1hr |
| Costs | $130-1200 | <$30 |
| Quality | 10 preferred | 10  preferred |

**MT + PE  saves time & costs and maintains quality**

# PHAST:
# Collaborative Translation Tool

Shared, collaborative web-based tool designed to facilitate translation of health promotion materials across health departments and programs.



**Public Health Automated System for Translation (PHAST)™**

PHAST is a collaborative translation management system for the public health domain developed at the University of Washington. It is currently in its beta version; access is by invitation only.

# Conclusions

- Tremendous need for language appropriate health materials

- Cost, time and expertise are barriers

- MT with public health post-editors can produce translated materials of equivalent quality in less time and therefore less costs

- Collaborative approach could provide way to implement quality MT system

# Next Steps

- Chinese (traditional)
- Test collaborative tool with a small number of health departments
- Expand dissemination

# Acknowledgments

- Katrin Kirchhoff, PhD – Co-PI, Dept. of Electrical Engineering
- Kristin Dew, Megumu Brownstein - Project Coordinators
- Students: Daniel Capurro, Kate Cole, Nathalie Martin, Loma Desai
- Staff of participating health departments
- NIH  National Library of Medicine

# Publications

- "Statistical Machine Translation of Public Health Information: A Feasibility Study." Kirchhoff, K., Turner, A.M., Axelrod, A., Saavedra, F. *JAMIA (2011).*

- "Evaluating User Preferences in Machine Translation Using Conjoint Analysis." Kirchhoff, K., Capurro, D. and Turner, A.M *, Proceedings of Association for Computer Linguistics (2012).*

- "Using Crowdsourcing Technology for Testing Multilingual Public Health Promotion Materials." Turner, A.M., Kirchhoff, K. and Capurro, D. *JMIR 14:12 (2012).*

- "A comparison of human and machine translation for public health practice: time, costs and quality." Turner, A.M., Bergman, M., Brownstein, M., Cole, K. and Kirchhoff, K. *J Public Health Manag Pract* (2013).

- "Modeling translation workflow to design machine translation applications for public health practice." Turner, A.M., Brownstein, M., Cole, K. , Karasz, H. and Kirchhoff, K. *J of Biomed Inform* (2014).

# The end

- Contact:

Anne M. Turner

[amturner@uw.edu](mailto:amturner@uw.edu)