

Developing a highly specialized corpus of spoken English: AAC discourse in the workplace

Carrie Bruce, Eric Friginal, Pamela Pearson
& Lucy Pickering

AAACL 2009: University of Alberta, Alberta, Canada

Session overview

- ◉ Introduction
 - > The challenge & our project
- ◉ Method
 - > Data collection & issues
 - > The AAC Workplace Discourse Corpus
 - > Analysis
- ◉ Future Directions
 - > Where do we go from here?

Introduction

- ◉ Collaborative project
 - › Georgia State University Department of AL (Lucy Pickering, Eric Friginal & Pam Pearson)
 - &
 - › Georgia Institute of Technology Center for Assistive Technology & Environmental Access (Carrie Bruce)
- ◉ What is AAC?
 - › Augmentative and assistive communication
 - › Example AAC device and user

Introduction

- ◎ 3,500,000+ individuals
 - > with severe communication impairment in the US alone (Beukelman & Mirenda, 2005)
- ◎ AAC devices are presumed to
 - > facilitate interaction, and
 - > promote fuller participation in social situations (e.g., school or work)

Introduction

- However...
 - AAC devices have been found ineffective at providing quick access to the context-specific language needed in work settings (Bryen, Potts, & Carey, 2007)
 - AAC users frequently report that the existing pre-programmed “work” pages do not facilitate their workplace tasks (Bruce, personal communication)

Our project

◉ Aims to

- 1) Examine the lexicogrammatical features & discourse patterns used by AAC- and non-AAC users in similar jobs
- 2) Address current inadequacies of AAC devices in the workplace by providing a more comprehensive and rigorous understanding of the language needed for workplace tasks & interaction
- 3) Guide the creation of pre-programmed pages and the NLP techniques used in AAC prediction systems

Method

- Data collection mirrored Holmes (2000) *Language in the Workplace Project*
 - > Speech samples were
 - (a) gathered over five consecutive work days to ensure a wide range of routine and novel topics;
 - (b) captured via wearable digital voice recorders; &
 - (c) transcribed following an enhanced orthographic transcription scheme (Crowdy, 1994; BNC)

Data collection issues

○ Participant recruitment

- › AAC users (n = 4) & non-AAC users (n = 4)
- › However, recruitment proved difficult
 - Few AAC users work in jobs where interaction takes place; sensitive context (tax preparation, attorney's office, etc.); lack of supervisor support

○ Participant communication

- › E-mailing unreliable
- › Telephoning not possible

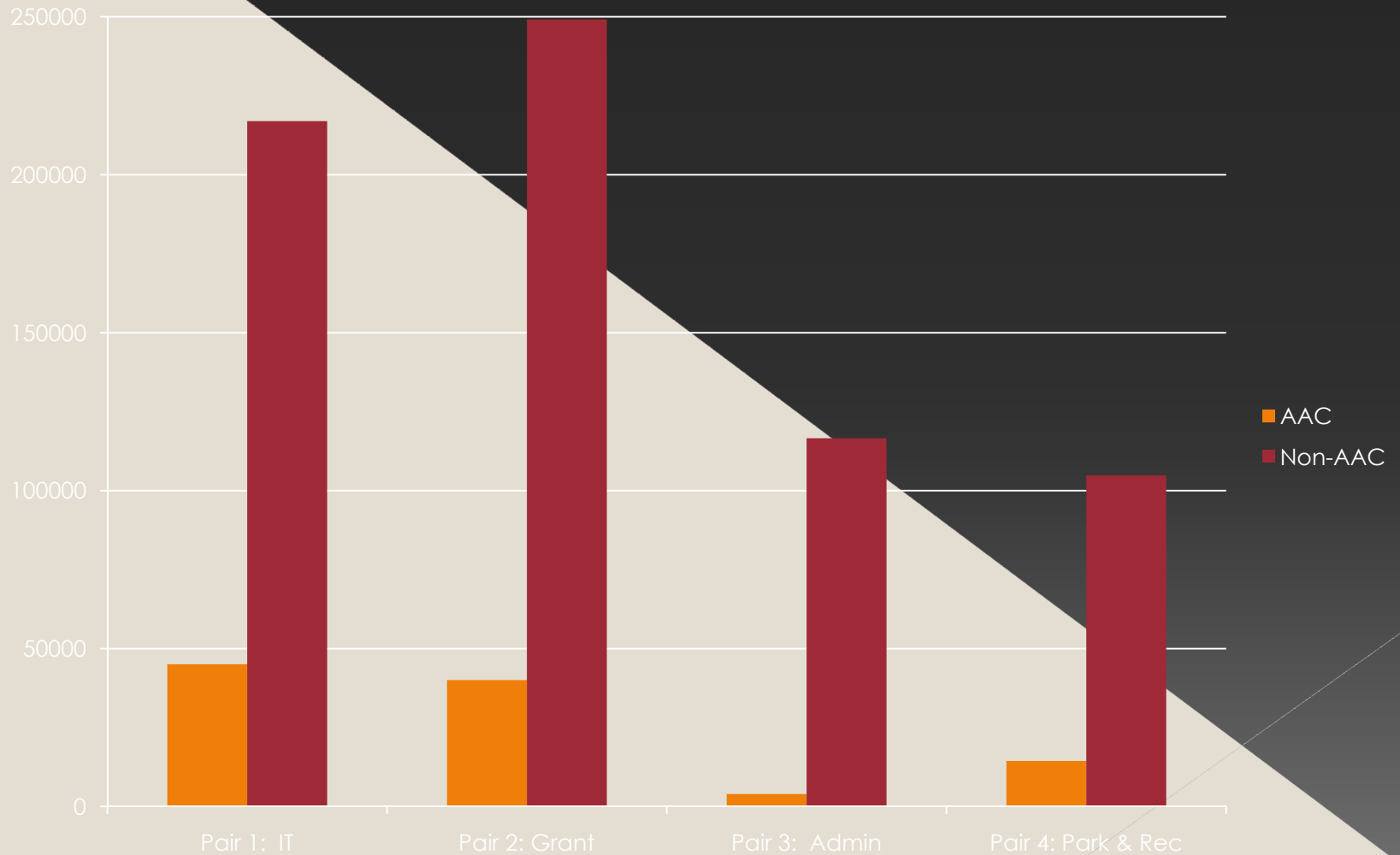
Data transcription issues

- ◉ Interaction vs. corpus-based views
 - > Importance of measuring pausing, noting speaker IDs
- ◉ Unique language situations
 - > [voc]
- ◉ Transcription team
 - > 12 trained UG transcribers
 - > Reliability checks for accuracy and transcription conventions

The corpus

- The AAC Workplace Discourse Corpus (Bruce & Pickering, 2009)
 - > 180+ hours
 - > Eight focal participants
 - > 60+ interlocutors
 - > Seven workplace settings
 - > Nearly 800,000 words in all
 - AAC \approx 103 K
 - Non-AAC \approx 688 K

Sub-corpora (by context & pair)



Data analysis

- Transcripts were cleaned and tagged
 - › Biber POS tagger (Biber, 2006)
- Currently working on:
 - › Lexicogrammatical & content analysis
 - Frequencies, topics, interactional patterns
 - › Comparison across corpora
 - Includes a measure of comparison with NZ *Workplace Project* (Holmes, 2005)
 - › MD analysis (Biber, 1988, 1995)

Future directions

- ◉ Video in addition to audio?
- ◉ Ultimate goals:
 - › To use findings to inform the pre-programming of AAC device pages for specific workplace contexts, and
 - › To test usability with AAC users

Conclusion

- We have a long way to go on this project...
 - But we feel that facilitating the workplace interaction of even a handful of AAC users is worthwhile
- Eventually, we hope to make a case for context-sensitive work pages to AAC manufacturers!

Questions?

- ◉ Thank you for listening to our paper!
- ◉ Contact information
 - › Carrie Bruce (AAC devices, NLP & page pre-programming): carrie.bruce@coa.gatech.edu
 - › Eric Friginal (spoken corpus Qs): eslef@langate.gsu.edu
 - › Pam Pearson (ppt): pampear@gsu.edu
 - › Lucy Pickering (general Qs): esllup@langate.gsu.edu