

Xenophenomena: studies of foreign language influence at several linguistic levels

Anders Lindström^{†‡} and Robert Eklund^{†‡}

anders@xenophones.com robert@xenophones.com

 Teli Research AB, Farsta



[‡] Dept. of Information and Computer Science, NLPLAB, Linköping University

Telia Research AB

1 DGIS 2002, Feb. 27



Outline

- Background
- Xenophones, a production study
- Xenomorphology
- Related work in the Speech Tech. Community

Telia Research AB

2 DGIS 2002, Feb. 27



Background

1995

- Speech data collection for Swedish ASR (Robert)
- Name pronunciation for Swedish TTS (Anders)
- Ambition to approach this "the proper way"

1996-2001

- A number of publications

Telia Research AB

3 DGIS 2002, Feb. 27



Bibliography (<http://www.xenophones.com>, <http://xenophones.info>)

Robert Eklund & Anders Lindström. 2001.

Xenophones: An Investigation of Phone Set Expansion in Swedish and Implications for Speech Recognition and Speech Synthesis.
Speech Communication 35 (2001), pp. 81–102.

Anders Lindström & Anna Kasaty. 2000.

A two-level approach to the handling of foreign items in Swedish speech technology applications.
Proceedings of ICSLP'00, Beijing, 16–20 October, vol. 1, pp. 54–57.

Anders Lindström & Robert Eklund. 2000.

How foreign are "foreign" speech sounds? Implications for speech recognition and speech synthesis.
Multi-Lingual Interoperability in Speech Technology, RTO Meeting Proceedings 28.

Papers and reports presented at the Tutorial and Workshop held in Leusden, The Netherlands, 13–14 September 1999.

Anders Lindström & Robert Eklund. 1999.

Xenophones Revisited: Linguistic and other underlying factors affecting the pronunciation of foreign items in Swedish.
Proceedings of ICPhS 99, San Francisco, California, 1–7 August 1999. Paper 0708, vol. 3, pp. 2227–2230.

Anders Lindström & Robert Eklund. 1999.

[j̥æ:mes] or [dʒæ:imz] or Perhaps Something In-between? Recapping Three Years of Xenophone Studies.
Gothenburg Papers in Theoretical Linguistics, 81.

Proceedings Fonetik 99, The Swedish Phonetics Conference, June 2–4 1999, pp. 109–112.

Robert Eklund & Anders Lindström. 1998.

How To Handle "Foreign" Sounds in Swedish Text-to-Speech Conversion: Approaching the 'Xenophone' Problem.
Proceedings of ICSLP 98, Sydney, November 30 – December 5. Paper 514, Vol. 7, pp. 2831–2834.
CD-ROM from Causal Productions Pty Ltd, PO Box 100, mailto:info@causal.on.net.

Robert Eklund & Anders Lindström. 1996.

Pronunciation in an internationalized society: a multi-dimensional problem considered.
FONETIK 96, Swedish Phonetics Conference, Nässlingen, 29–31 May, 1996.
TMH-QPSR 2/1996, pp. 123–126.

Telia Research AB

4 DGIS 2002, Feb. 27



Background

- ASR and TTS high-quality enough to be included in public services. Problems to be solved remain
- Increasingly globalized economy and society means spoken interfaces must handle foreign features on several linguistic levels
- Number of such applications grows with increasing mobility and tourism (Billi)

Applications

- Travel bookings, booking cinema tickets
- International news
- Navigation systems
- ... etc.

- Swedish user, using an English speech- and pen-based GPS navigational system in Poland...

Introduction

- A field linguist with no prior knowledge of Swedish makes recordings at a pub...
- The task: to describe the phonemic system

Swedish linguistic data*

- Vem i The Beatles gjorde den bästa solodebuten, tycker du?
- John Lennons platta blev ju mest betydande...
- Ja, fast George Harrison fick största hiten.
- Strunt i det. Ska vi se nya James Bond?
- Har inte sett Bond sen Timothy Dalton tog över rollen.
- Men nu är det ju Pierce Brosnan som spelar Bond.

* *No translation, you're a field linguist, remember!*

“Swedish” linguistic data?

- Vem i The Beatles gjorde den bästa solodebuten, tycker du?
- John Lennons platta blev ju mest betydande...
- Ja, fast George Harrison fick största hiten.
- Strunt i det. Ska vi se nya James Bond?
- Har inte sett Bond sen Timothy Dalton tog över rollen.
- Men nu är det ju Pierce Brosnan som spelar Bond.

A speech technology angle

- Swedish Text-to-Speech Systems and Automatic Speech Recognition Systems will come across these speech sounds
- What's the best choice for TTS?
- What variability will an ASR system encounter?

A phonological perspective

- What is the status of “foreign” speech sounds?
 - Phonemes (change of meaning?)
 - Allophones (of what?)
 - Marginal? (quite central in some contexts)
 - ... or simply: phones of foreign origin
- Introduced the term *xenophones* in 1998

What do language users do?

- Native Swedish speaker talking about British PMs
- Variation:

near the source language



something inbetween

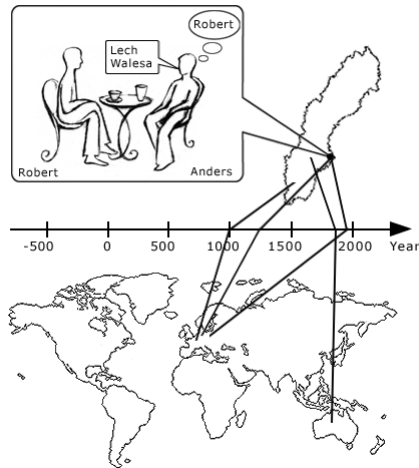


rephonematized

Research questions:

- What do language users do?
- ... and why?

Factors



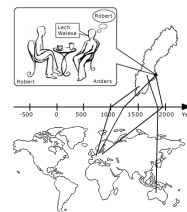
Telia Research AB

13 DGIS 2002, Feb. 27



Possible factors

- A's competence
- A's performance
- A's expectations on R's competence
- Social distance between A and R
- Socio-cultural distance to the source language
- Time, place and channel used when borrowing
- "Distance" between the phonological systems involved
- A's regional background
- et cetera



Telia Research AB

14 DGIS 2002, Feb. 27



Method. Production study

- Speech sounds not included in traditional descriptions of Swedish
- Phones of predominantly Anglosaxon origin
[tʃ dʒ ʃ ʒ θ ð z w aɪ eɪ əʊ juː æ]
- Phones of Germanic origin
[x aː]

Sentences (examples)

Många har Roger Moore som favorit i rollen som James Bond
Many prefer Roger Moore's interpretation of James Bond

Veckopressens favoriter är verkligen Diana och Charles
Diana and Charles are indeed the favourites of the tabloids

Intercity-tåget gick direkt från Aachen till Baden-Baden
The Intercity train went straight from Aachen to Baden-Baden

Det anses allmänt att John Major är en blek efterträdare till Maggie Thatcher
It is a widely held opinion that John Major is a rather pale successor to Maggie Thatcher

Recordings

- 40 locations
- 490 people, 15-75 years old, Telia employees (or their relatives)
- 12 sentences among 5400 (per person), constituting 1-2 minutes in a one-hour session
- 23 750 xenophone tokens

Evaluation

- Narrow phonetic transcription by four Swedish transcribers with above-the-average knowledge of English
- Did not want native speakers of English: Focus on what sounds non-Swedish to native speakers of Swedish





Examples

<i>James</i>	[ja:mes]	Rephonematized
	[jejms]	} "In-between"
	[jejmz]	
	[dʒɛɪms]	Near source-language
	[dʒɛɪmz]	

Variability

Det anses allmänt att John Major är en blek efterträdare till Maggie Thatcher

It is a widely held opinion that John Major is a rather pale successor to Maggie Thatcher

1. Close to the source language 
2. Some accommodation to Swedish 
3. Rephonematization 
4. Problems! 

Results

- Less than 3% resort to total rephonematization
- Most expand their phone repertoire considerably
- Variation with respect to
 - Xenophone
 - Word/name
 - Position i word/name

Tables

- Along two dimensions: awareness and fidelity

	Aware	Not aware
High fidelity	Category 1: succeed in producing xenophones	—
Low fidelity	Category 2: try to produce something non- Swedish	Category 3: rephonemati- zation

Example table

			1	2	3
Awareness			High	High	Low (or high?)
"Fidelity"			High	Low	Low
Xeno- phone	No.	In the word			
θ	456	Thriller	49.6 θ	0.2 ç	48.0 t
					2.2 t
θ	463	Thatcher	42.3 θ	1.3 tç	56.0 t
				0.2 s	
				0.2 *	

Telia Research AB

23 DGIS 2002, Feb. 27



Results (consonants)

- Affricates common:
 - [tʃ] > 90%
 - [dʒ] > 20–60% (depending on word/name)
- Dental fricatives produced by approx. 50%:
 - [θ ð]
- Rarely any voiced fricatives at all (< 1%):
 - [ʒ z]

Telia Research AB

24 DGIS 2002, Feb. 27



Results (vowels)

- All vowels/diphthongs well approximated (approx. 90%). [aɪ eɪ əʊ juː æ]
- [aɪ eɪ] well approximated by Swedish /V+j/.
- Perhaps surprising: [əʊ]

Results

- Gender:
 - Small differences, only one weakly significant
- Region
 - Regions based on Elert's dialectal areas
 - Differences exist but hard to find patterns

Results: Age

- Plays a major role
- Seven decadic intervals with boundaries at 15, 25, 35, 45, 55, 65, 75 years
- Significant differences between all pairs of groups, except the youngest

Results: Education

- Four categories
 - "folkskola" (approx. 7 years)
 - "realskola/grundskola" (approx. 9 years)
 - "gymnasium" (11-13 years)
 - University education
- Clear differences: higher education yields more xenophones (cf. Ljung)
 - Nota bene! His subjects were aware of the objective of the study.

Implications for ASR

(ASR = Automatic Speech Recognition)

- More units/less specific acoustic models
 - Not clear how this fits with current state-of-the-art technology (HMMs)
- Bilingual recognizers?
- Domain and application dependent

Implications for TTS



- What is the "proper" level?
 - Rephonematized [jɑ:mes]
 - "Too good" synthesis: [dʒɛɪmz]
 - "Swedish" pronunciation: [dʒɛɪms]

Synthesis examples

- Xenophones included in our synthesis recordings since 1996
- Möbius (Bell Labs), 1997
- Example:

Det anses allmänt att John Major är en blek efterträdare till Maggie Thatcher

It is a widely held opinion that John Major is a rather pale successor to Maggie Thatcher

1. Without xenophones 
2. With xenophones 

Telia Research AB

31 DGIS 2002, Feb. 27



Xenomorphology



- 415 strips from the comic *Rocky* by Martin Kellerman, who claims to be writing down what his friends say.
- Anglosaxon dominance, some German and French items. Derivation, inflection and compounding.

Telia Research AB

32 DGIS 2002, Feb. 27



Lexical implementation

- Lindström & Kasaty (2000)
- Users good at assigning inflectional paradigms to nonsense words (Linell, Granberg)
- Two-level formalism (Antworth's PC-Kimmo)
- Handling constraints on morph-level features:
 - morphotactics
 - cross-lingual combinability
 - word prosody

A model for...

- Foreign items in Swedish
 - on the phonetic, phonological (including prosodic), morphological, possibly also syntactic level
 - in a synchronous perspective
 - which holds for observed data
 - and can explain the data in terms of underlying factors
- Lexical implementation (Lindström & Kasaty)

Related speech technology research

- *Trancoso, Fitt, Ladefoged & Maddieson, Möbius, Steeneken, Hunt & Mandel, Johnston, Black & Font-Llitjós, Lüdeling et al., Nass & Dahlbäck, Oflazer*

Sweden

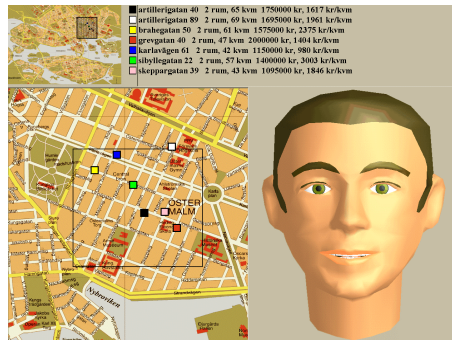
- *Ljung (EIS), Sharp (PhD, 2001)*
- *Chen & Sääv (Eurospeech 2001)*
- *Beskontjina (English in Russian, 2001)*

Related speech technology conferences

- *Eurospeech/Interspeech, ICSLP. ISCA*
- *NATO-MIST (Leusden, 1999)*
- *Workshop on Multilingual Speech and Language Processing 2001 (Aalborg, 2001)*
- *Pronunciation Modeling and Lexicon Adaptation for Spoken Language Technology, (Colorado, 14-15 Sep. 2002)*

Related research (cont.)

- Telia Research and CTT, Royal Inst. of Technology: AdApt (Bell, Boye, Gustafson, Wirén)



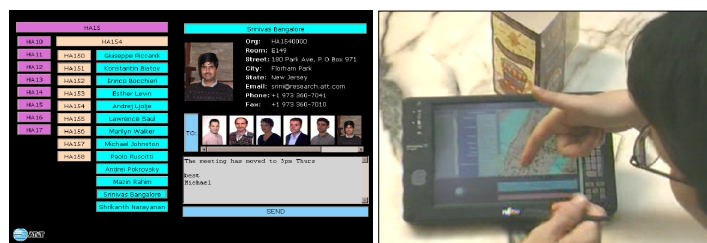
Telia Research AB

37 DGIS 2002, Feb. 27



Related research (cont.)

- AT&T
 - PDAMVPQ: Multimodal access to corporate directory and messaging
 - MATCH - Multimodal Access To City Help



Telia Research AB

38 DGIS 2002, Feb. 27



Related research (cont.)

- Lucent and e.Digital: *VoiceNav™ voice navigation*
- MXP 100:



Discussion

- Xenophones need to be handled in ASR/TTS. How?
- The production study gave some insight in attitudes to foreign speech sounds, but perceptions studies should be made
- Factors? Are we studying language or etymological knowledge, knowledge of concepts/names/people or general knowledge of the world?
- The role of regional background unclear
- Cultural differences? Other source languages?