

SVOX MOBILE: MEETING QUALITY EXPECTATIONS IN OVER 30 LANGUAGES

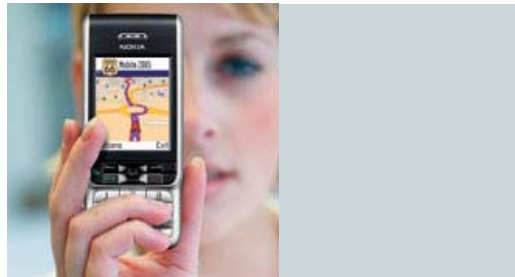
Nikos Chatzichrisafis
Project Manager



Voice Search 2008

SVOX Text-To-Speech Product Family

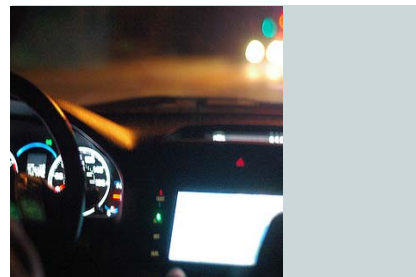
Mobile Devices



- ▶ Mobile navigation devices
- ▶ PDAs
- ▶ Mobile phones

- ▶ Navigation
- ▶ Location based services
- ▶ Personalization

Automotive



- ▶ Navigation systems
- ▶ Multimedia systems
- ▶ Hands-free car kits

- ▶ Voice guidance
- ▶ Messaging
- ▶ News-on-demand

Service Providers



- ▶ Datacenter Servers

- ▶ Interactive Voice Response
- ▶ Location Based Services
- ▶ ...

Platform

Applications

Quality Expectations

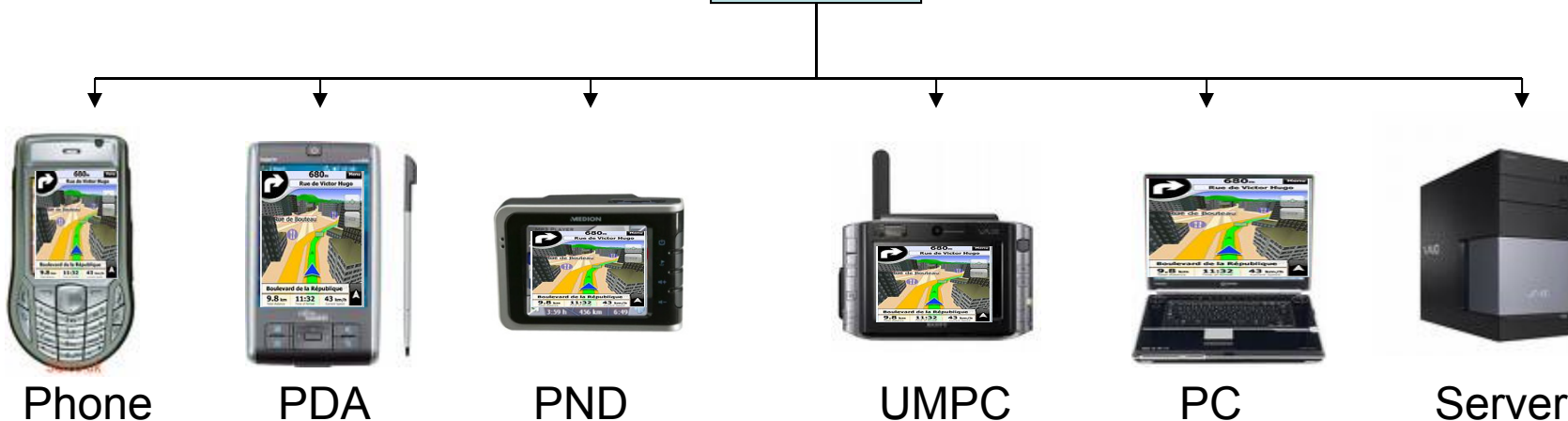
- System Developers
 - Easy integration
 - Computational efficiency
 - Flexibility
- End Users
 - Intuitive interfaces

Easy Integration

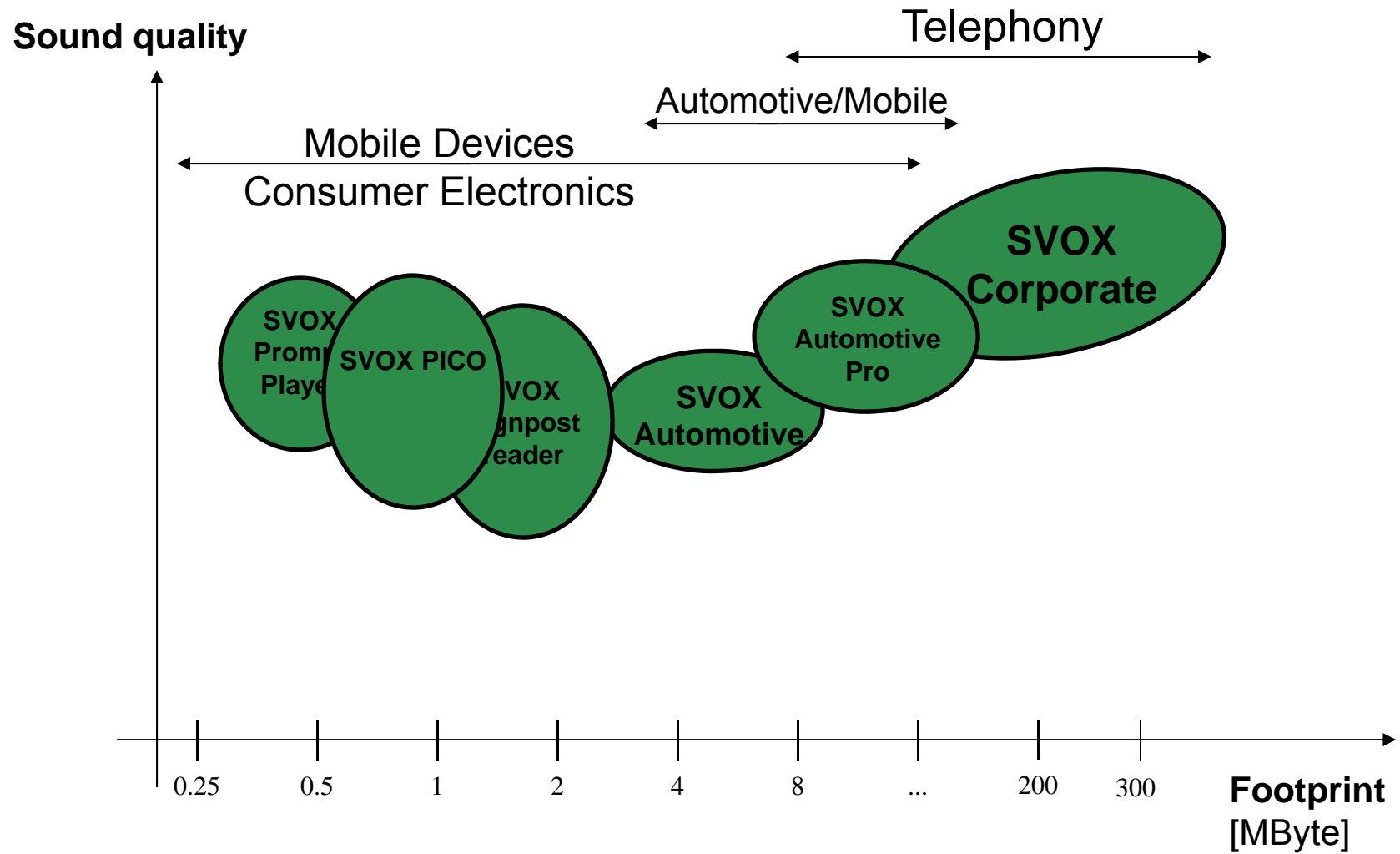
- Same API for all platforms
- Lingware Voice Package tailored to device requirements



Application with integrated SVOX Engine



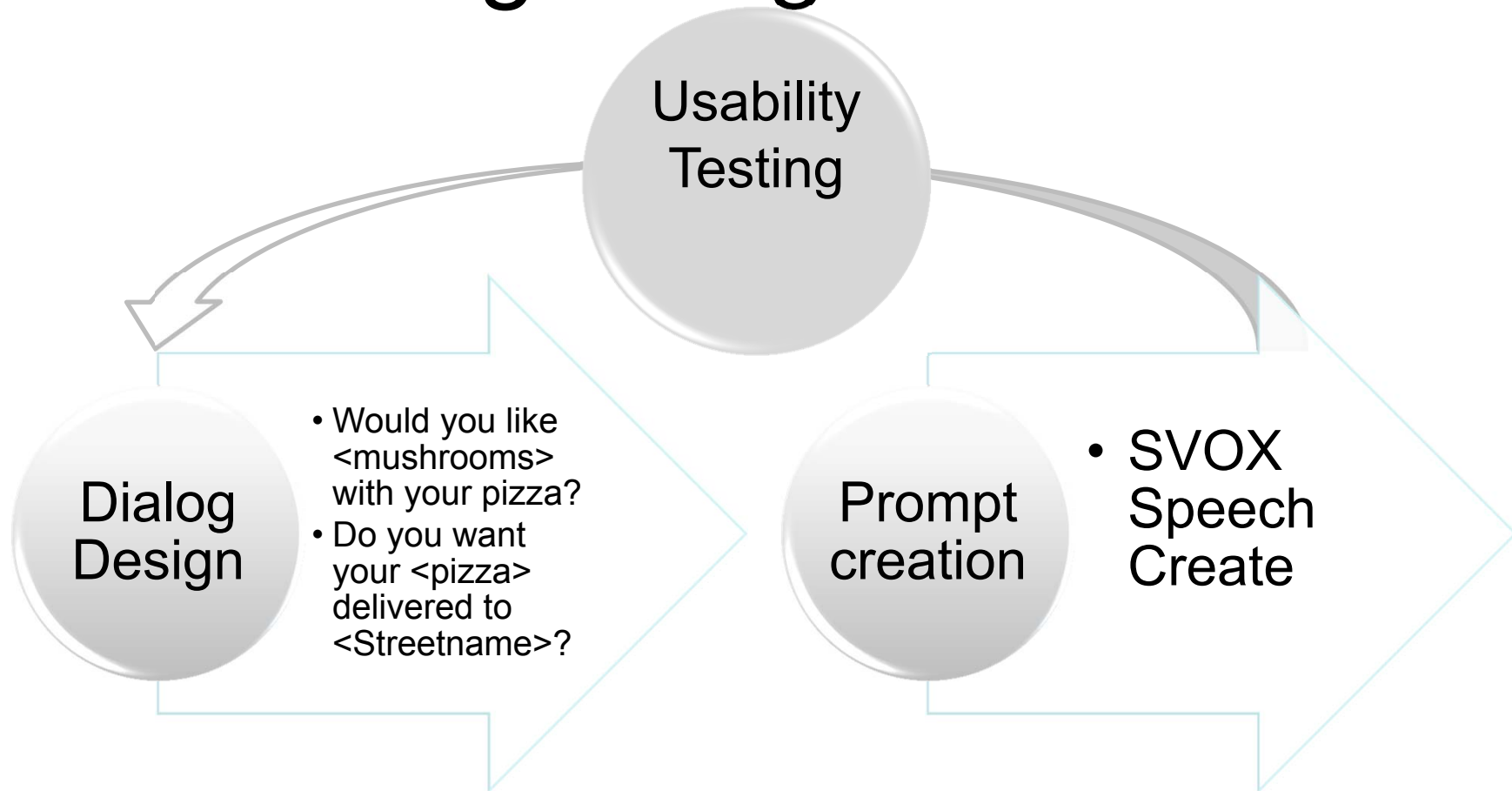
Scalable Lingware



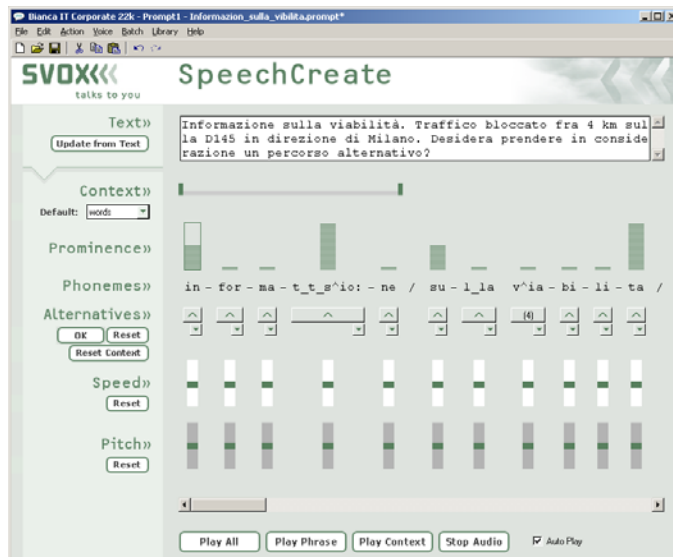
Quality Expectations

- System Developers
 - Easy integration ✓
 - Computational efficiency ✓
 - Flexibility
 - Dialog Design
 - Language Deployments
- End Users
 - Intuitive interfaces

Dialog Design Solution





SVOX SpeechCreate



- Controllable speech output

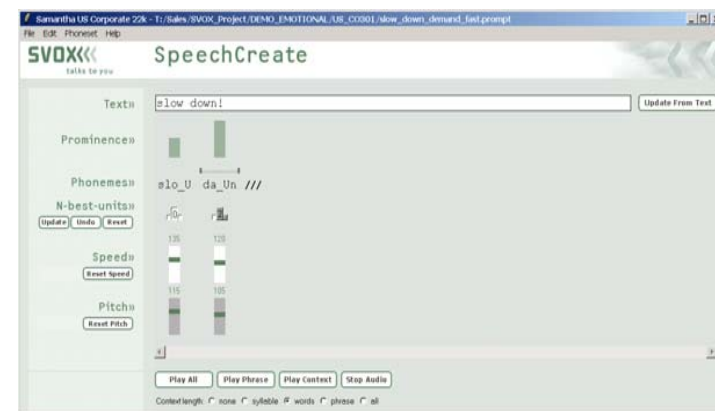





Prosody modification

- Take the first exit on your left. 
- Take the **first** exit on your left. 

SpeechCreate: Urgency

- Designing the content of the message, e.g. from traffic-congestion forecast:
 - "slow down!"
- Use SVOX SpeechCreate to specify:
 - level of urgency, emotion
 - Pitch, speed
- Optimize prompts to suit the demands of the driving situation to ensure a better driver response time



neutral	urgent	urgent-fast
		
0.88 s	0.83 s	0.70 s

Problems with voice talents

Lead Time

- Lead time for additional recordings is typically in the order of weeks
- Desired lead time shortly before production is in the order of hours or days

Availability

- Professional speakers are not always available
- Speakers can become completely unavailable. In this case, all recordings have to be done from scratch

Quality

- The sound of a voice changes over time and from day to day – old and new recordings do not match well

Cost

Quality Expectations

- System Developers
 - Easy integration ✓
 - Computational efficiency ✓
 - Flexibility
 - Dialog design ✓
 - Language Deployments
- End Users
 - Intuitive interfaces
 - Customization

“Preparing for Localization”



– To achieve an acceptable level of localization, follow these basic steps:

1. Basic Translation
2. Nationalization
3. Regionalization
4. Expatriates
5. Customization for Corporations and Industries
6. Forms of Addresses
7. Male or Female

Localization pitfalls (1/3)

- English
 - “turn left, Sequoia Drive.”
 - “turn left onto Sequoia Drive.”

Localization pitfalls (1/3)

- English
 - “turn left, Sequoia Drive.”
 - “turn left onto Sequoia Drive.”
- Spanish
 - gire a la derecha, Avenida de Andalucía
 - **la** Avenida de Andalucía
 - **e**l Camino Real
 - **e**l Carril Principal

Localization pitfalls (2/3)

- French

5 MB system!

- ▶ Dans 800m prendre la 2 sortie vers l'avenue Victor Hugo
- ▶ Dans 800m prendre la 2 sortie vers le rue del'Escargot.



- German

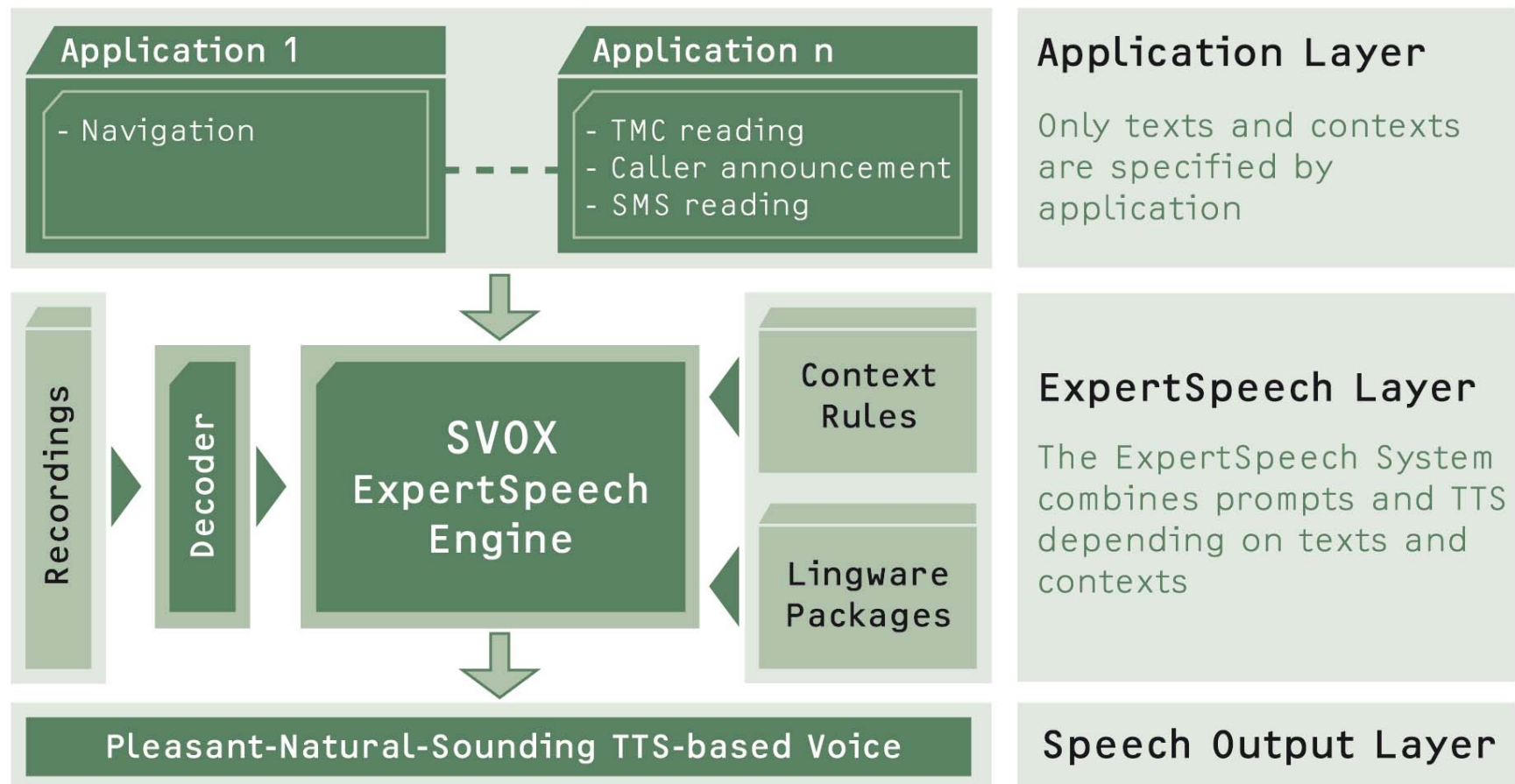
- ▶ biegen sie links ab in die HAUPTSTRASSE.
- ▶ biegen sie links ab in den BURGGRABEN.
- ▶ biegen sie links ab auf den HESSE-PFAD.
- ▶ biegen sie links ab in den HESSE-TUNNEL.
- ▶ bleiben sie links auf der HAUPTSTRASSE.
- ▶ bleiben sie links auf dem HILLGARDGÄSSCHEN.



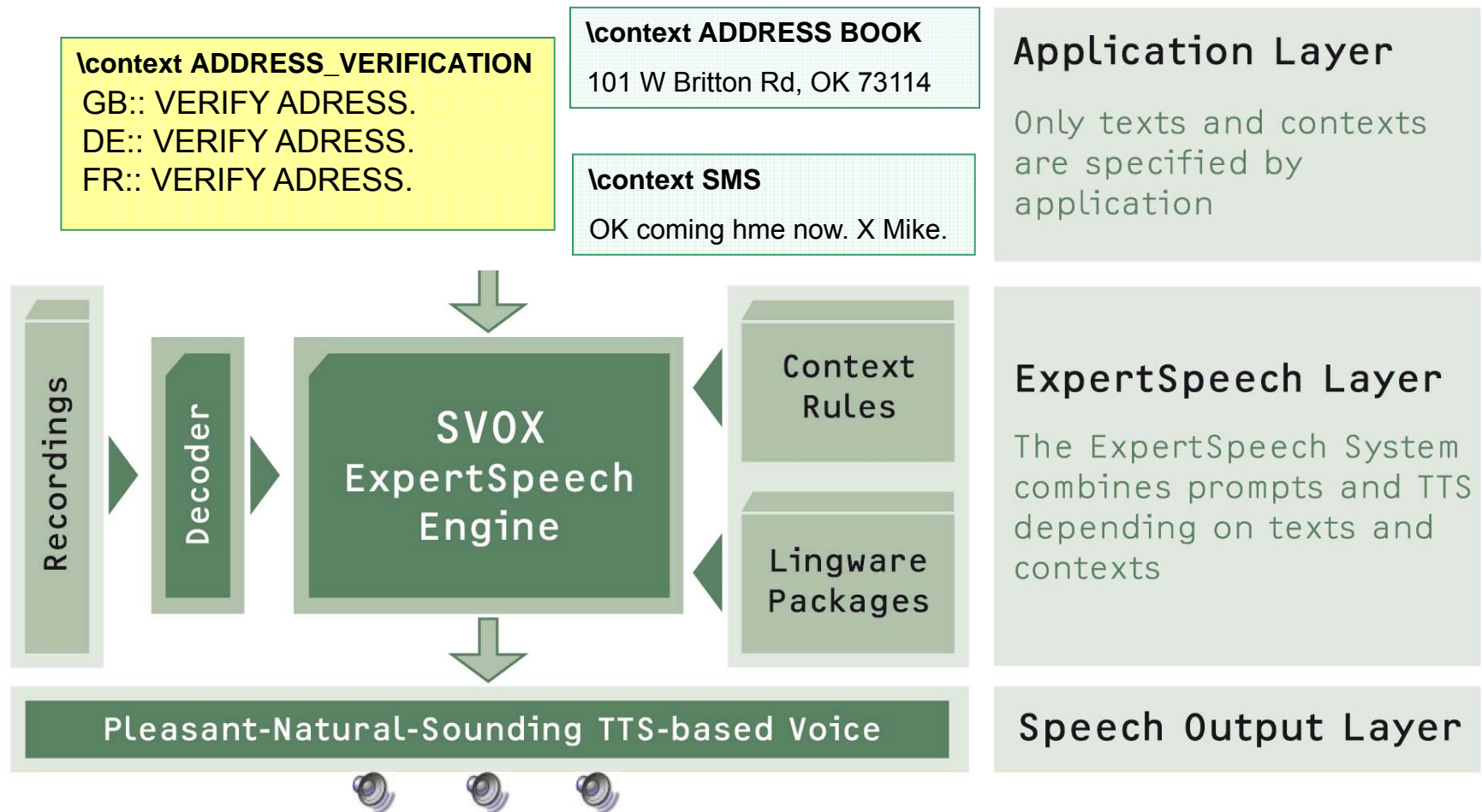
Localization pitfalls (3/3)

- Swedish
 - Är "Våxnäsgatan 21" korrekt?
 - tjugo-ett, tjugo-en
- Greek
 - Έχετε παραγγείλει 1 αναψυκτικό και 1 πίτσα.
 - Ένα, μία

Expert Speech System



Expert Speech for different applications



Quality Expectations

- System Developers
 - Easy integration ✓
 - Computational efficiency ✓
 - Flexibility
 - Dialog design ✓
 - Language Deployments ✓
- End Users
 - Intuitive interfaces ✓

Mixed-Linguality

Place and Street Names for **Signpost Reading**

Major place names often have an **expected pronunciation** depending on the language:

e.g. Paris(EN)  Paris(FR)  Paris(DE) 

e.g. Brussels(EN)  Bruxelles(FR)  Brüssel(DE) 



For other place names, SVOX TTS maps the pronunciation to the **HMI language** using **the same voice**:

e.g. a Dutch voice saying „Gravelines“ (French town)

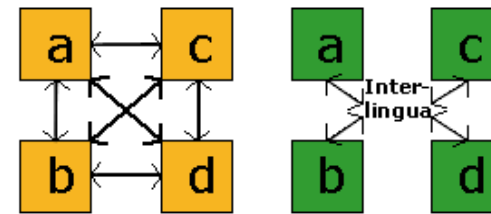
e.g. A French voice saying „Borssele“ (Dutch town)

No linguistic knowledge	With linguistic knowledge	reference
-------------------------	---------------------------	-----------



Mixed-linguality

Flexible, interlingual approach to phoneme-mapping across languages:



1. Standard Transfer Component 2. Interlingua

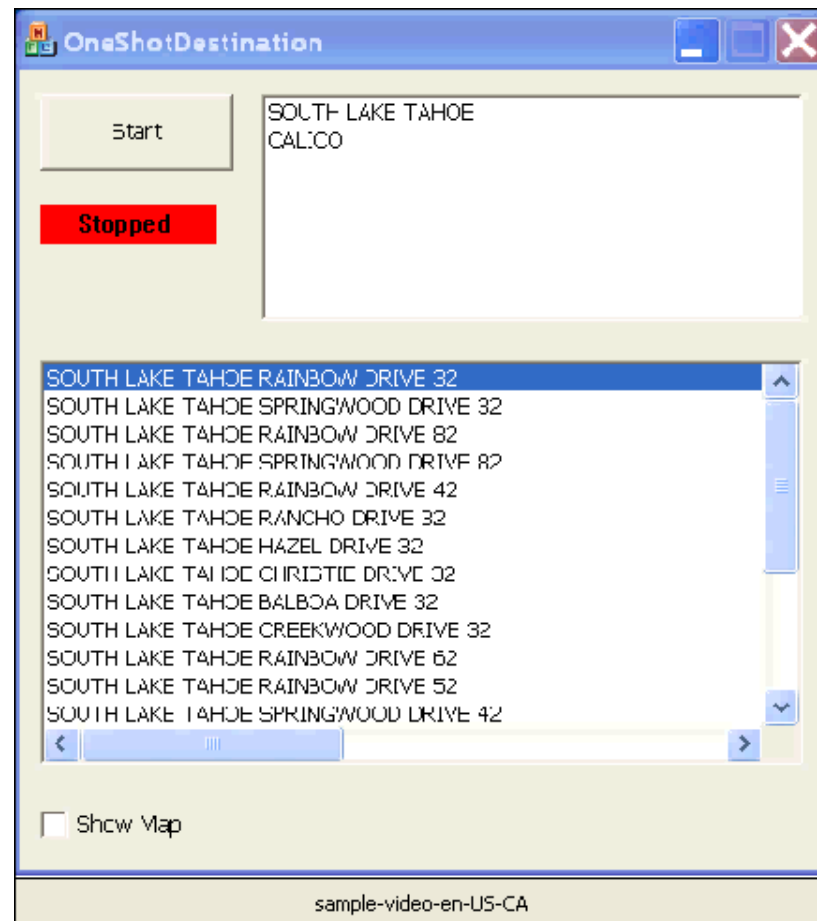
	Australia	Belgium	US	France	Germany	Ireland	Italy	Nether-lands	Portugal	Sweden
en-UK										
ru-RU										
de-DE										
es-MX										
ca-FR										
nl-NL										
sv-SE										

Quality Expectations

- System Developers
 - Easy integration ✓
 - Computational efficiency ✓
 - Flexibility
 - Dialog design ✓
 - Language Deployments ✓ ✓
- End Users
 - Intuitive interfaces ✓ ✓

VoiceSearch Demo (California)

- Destination Entry
- All Californian cities and streets



Quality Expectations

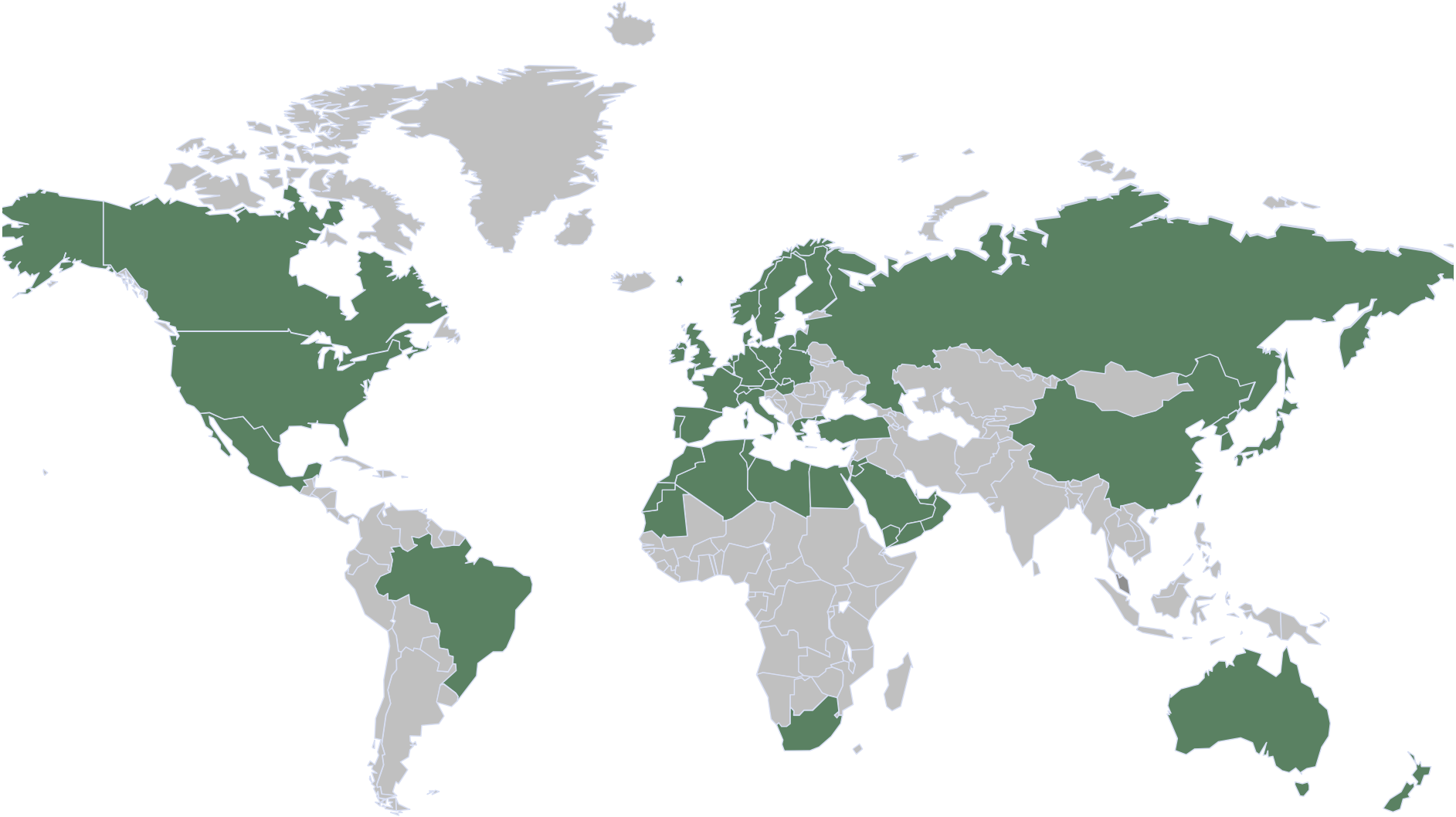
- System Developers
 - Easy integration ✓
 - Computational efficiency ✓
 - Flexibility
 - Dialog design ✓✓
 - Language Deployments ✓✓✓
- End Users
 - Intuitive interfaces ✓✓✓

Thank you!



Visit us at <http://svox.com/>

Language Coverage



SVOX Language Map