



VoiceObjects



Developer Jam Session

October 8, 2008



Multi-Language Self-Service

Best Practices for Design & Deployment of Multi-Lingual Dialogs

Dr. Andreas Volmer

avolmer@voiceobjects.com

Presales Manager EMEA



Dr. Angelika Salmen

asalmen@voiceobjects.com

Manager VUI Services CE





Agenda

- Managing multilingual prompts
 - Audio resource management best practices
 - Adding new languages with Storyboard Manager
 - Accounting for cultural and syntactic differences
- Let the caller choose
 - Initial / in-session language selection
- Natural pronunciation of dynamic data
 - The Formatting Bus (aka TTA Bus)



Analysis and Reporting

- Deployment Scenarios
 - Multi-language / multi-country deployments



Agenda

- Managing multilingual prompts
 - Audio resource management best practices
 - Adding new languages with Storyboard Manager
 - Accounting for cultural and syntactic differences
- Let the caller choose
 - Initial / in-session language selection
- Natural pronunciation of dynamic data
 - The Formatting Bus (aka TTA Bus)

Analysis and Reporting

- Deployment Scenarios
 - Multi-language / multi-country deployments



VoiceObjects Desktop

Managing multi-language audio prompts


Definition (...)


Name:


Audio Resource (...)

Specify an external audio file.

Location:






File: 

Extension: 

Random variations: 

Alternative Texts (...)

Specify the alternative texts which will be played back with TTS if the audio file is not available or if the output mode is TTS:Audio.

Alternative Text		English (UK) (1)
Alternative Text		English (US) (1)
Alternative Text		German (1)
Alternative Text		French (1)
Alternative Text		Italian (1)

In VoiceObjects Desktop, the designer controls when and how often a prompt is used in the call flow

VoiceObjects Desktop

Managing multi-language audio prompts

Definition (...)

Name:

Audio Resource (...)

Specify an external audio file.

Location:


File:

Extension:

Random variations:

Alternative Texts (...)

Specify the alternative texts which will be played back with TTS if the audio file is not available or if the output mode is TTS:Audio.

Alternative Text  English (UK) (1)

Label:

Layer:

Language:

Input mode:

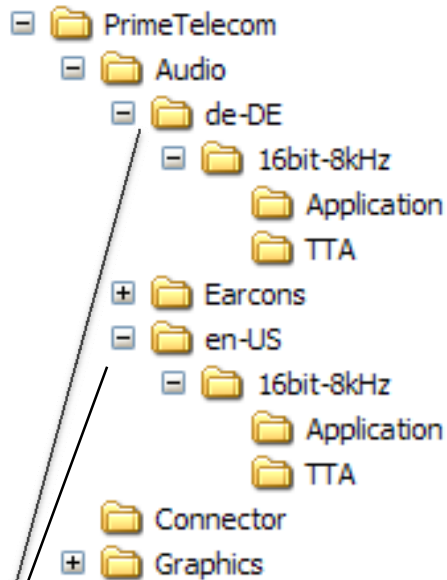
Variation:

Text:

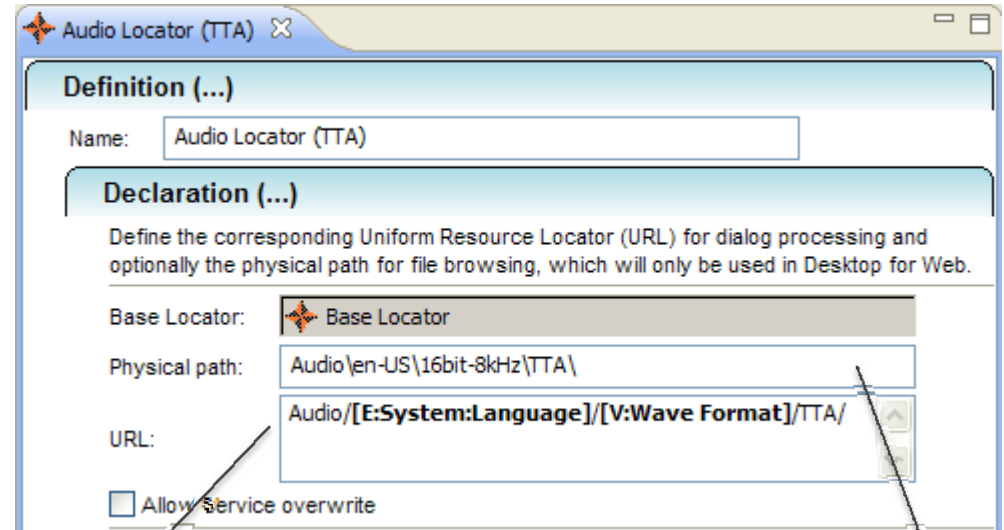
Dynamic Resource Locator



Using Dynamic Resource Locators



Separate base path per language, identical file names



Dynamic URL

Static Physical path (for file browsing in Audio editor)



Agenda

- Managing multilingual prompts
 - Audio resource management best practices
 - Adding new languages with Storyboard Manager
 - Accounting for cultural and syntactic differences
- Let the caller choose
 - Initial / in-session language selection
- Natural pronunciation of dynamic data
 - The Formatting Bus (aka TTA Bus)

Analysis and Reporting

- Deployment Scenarios
 - Multi-language / multi-country deployments



Storyboard Manager

Managing multi-language audio prompts



VoiceObjects

➔ Get Credit Card Type

	C	E	F	H	M	S	U	V	AD	AE
1	Module Name	Parent Object Type	Parent Object Name	Prompt Usage	Object Name	Audio File	Prompt Text	Language	Changes	Technical Notes
6	Enter new Credit Card	Input	Get Credit Card Type	Input Request	EditCreditCard	PT_EditCreditCard_1.wav	And what is the new credit card: Visa, MasterCard oder American Express?	English (UK)		
9	Enter new Credit Card	Input	Get Credit Card Type	Input Request	EditCreditCard	PT_EditCreditCard_1.wav	And what is the new credit card: Visa, MasterCard oder American Express?	English (US)		
10	Enter new Credit Card	Input	Get Credit Card Type	Input Request	EditCreditCard	PT_EditCreditCard_1.wav	Quelle carte de crédit avez-vous: Visa, MasterCard, où American Express?	French	NEW	
11	Enter new Credit Card	Input	Get Credit Card Type	Input Request	EditCreditCard	PT_EditCreditCard_1.wav				

Add Layer State

Add a new layer state on the basis of an existing prompt structure

Select the type of layer you want to add. Optionally, provide the desired state to use as a basis

Language
Base language: English (US)

Input mode
Base mode:

Custom layer
Base condition:

Provide a name for the new layer state

New state name: French

Copy prompts of base layer

Help OK Cancel

In Storyboard Manager, the VUI designer

- reviews and manages prompts
- adds prompt variations
- organizes recording sessions
- listens to prompt recordings
- creates new language layers

Storyboard Manager

Listening to prompts

1) Configure *Resource locator(s)* in *Resources* tab

	A	B
1	Resource Locator	URL
2	*Generic*	http://localhost:8070/VoiceObjects/Resources/
3	Audio Locator (Application)	http://localhost:8070/VoiceObjects/Resources/Samples/PrimeTelecom/Audio/en-US/16bit-8kHz/Application/
4	Audio Locator (Earcons)	http://localhost:8070/VoiceObjects/Resources/Samples/PrimeTelecom/Audio/Earcons/16bit-8kHz/

2) Double-click *Audio File* column in *Prompts* tab

	C	E	F	Q	S	U	V	Z
1	Module Name	Parent Object Type	Parent Object Name	Variation	Audio File	Prompt Text	Language	Resource Locator
15	Update Credit Card Expiration Date	Input	Get Credit Card Expiration date (web)	1	PT_CreditCardExpirationDateReprompt_1.wav	Enter the new expiration date of your credit card again, please.	English (US)	Audio Locator (Application)
16	Update Credit Card Expiration Date	Input	Get Credit Card Expiration date (web)	2	PT_CreditCardExpirationDateReprompt_2.wav	Please, enter the new expiration date of your credit card again.	English (US)	Audio Locator (Application)
17	Update Credit Card Expiration Date	Input	Get Credit Card Expiration date (web)	3	PT_CreditCardExpirationDateReprompt_3.wav	Enter the new expiration date of your credit card one more time, please.	English (US)	Audio Locator (Application)



Agenda

- **Managing multilingual prompts**
 - Audio resource management best practices
 - Adding new languages with Storyboard Manager
 - Accounting for cultural and syntactic differences
- Let the caller choose
 - Initial / in-session language selection
- Natural pronunciation of dynamic data
 - The Formatting Bus (aka TTA Bus)



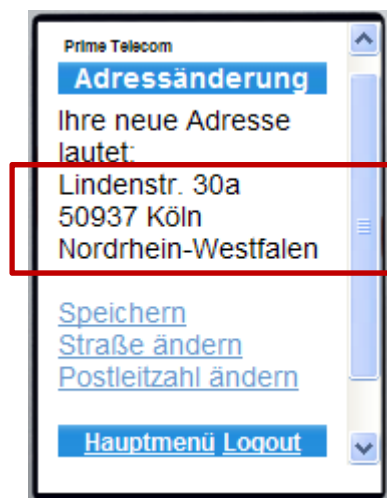
Analysis and Reporting

- **Deployment Scenarios**
 - Multi-language / multi-country deployments

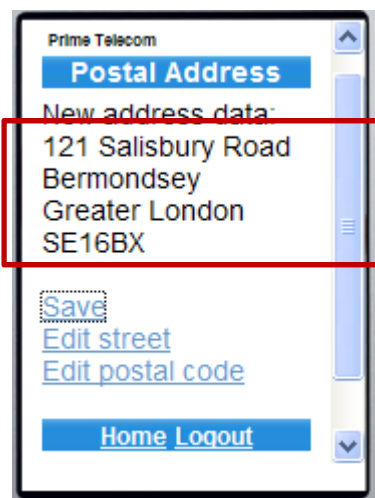


Accounting for cultural differences

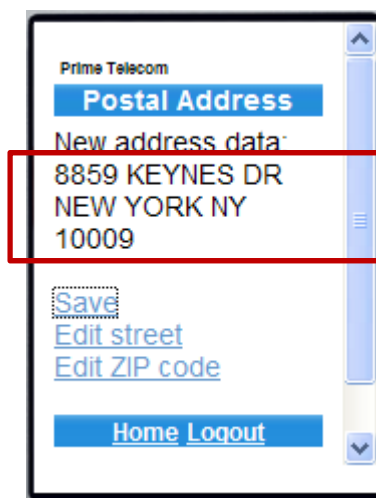
Example: Displaying a Postal Address



de-DE



en-UK

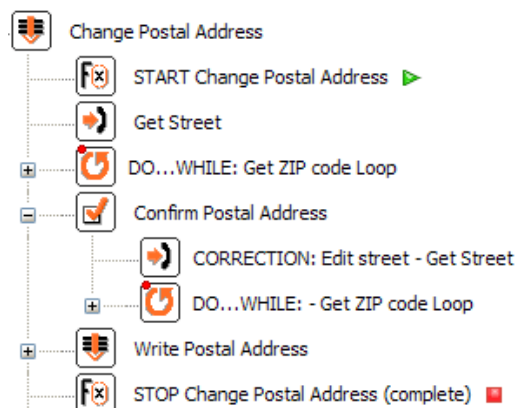


en-US



Accounting for cultural differences

Example: Displaying a Postal Address



Output Item	English (UK) (Always), Text/Web
Output Item	English (US) (Always), Text/Web
Output Item	German (Always), Text/Web

Label:

Layer:

Channel: Text/Web

Language: German

Occurrence: Always

Input mode: Default

Barge-in: Default

Prompt: [V:Address Street], [V:Address ZIP Code] [V:Address City], [V:Address County/State]

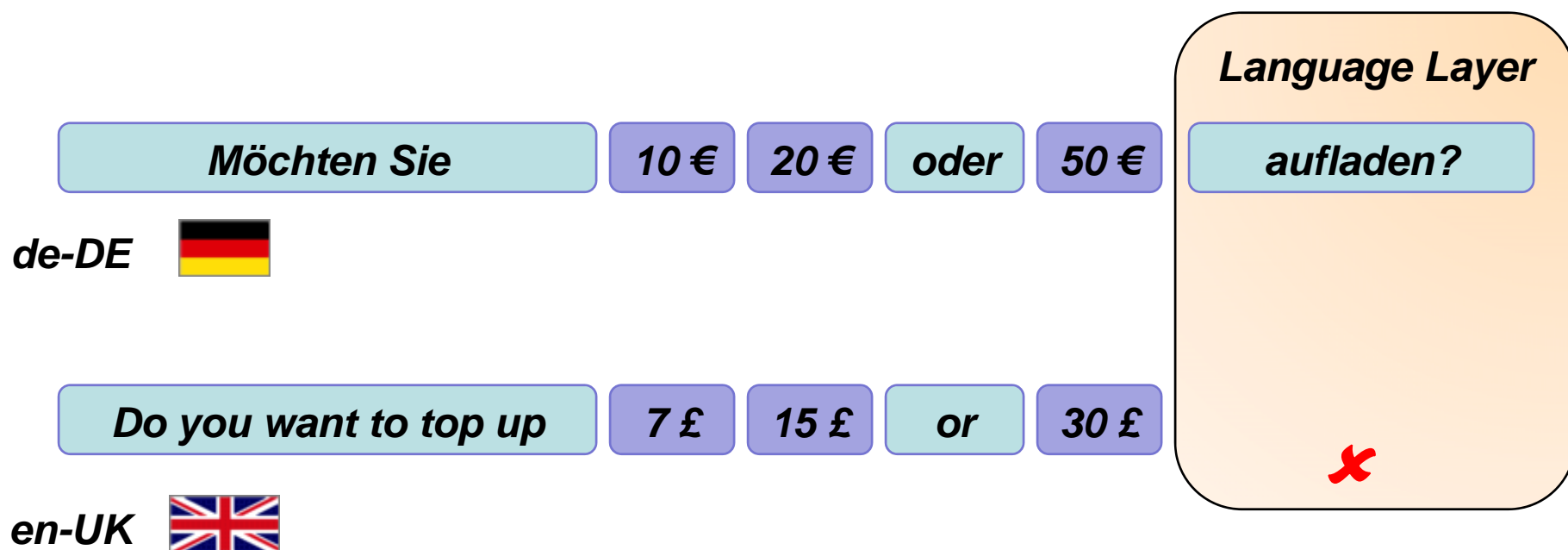
Lindenstr. 30a
50937 Köln
Nordrhein-Westfalen

• The Confirmation Request in **Confirm Postal Address** references the Output object **Format Address**

• Note usage of *Language* filter in Output Items



Accounting for syntactic differences





Agenda

- Managing multilingual prompts
 - Audio resource management best practices
 - Adding new languages with Storyboard Manager
 - Accounting for cultural and syntactic differences
- Let the caller choose
 - Initial / in-session language selection
- Natural pronunciation of dynamic data
 - The Formatting Bus (aka TTA Bus)

Analysis and Reporting

- Deployment Scenarios
 - Multi-language / multi-country deployments





Multi-language grammars

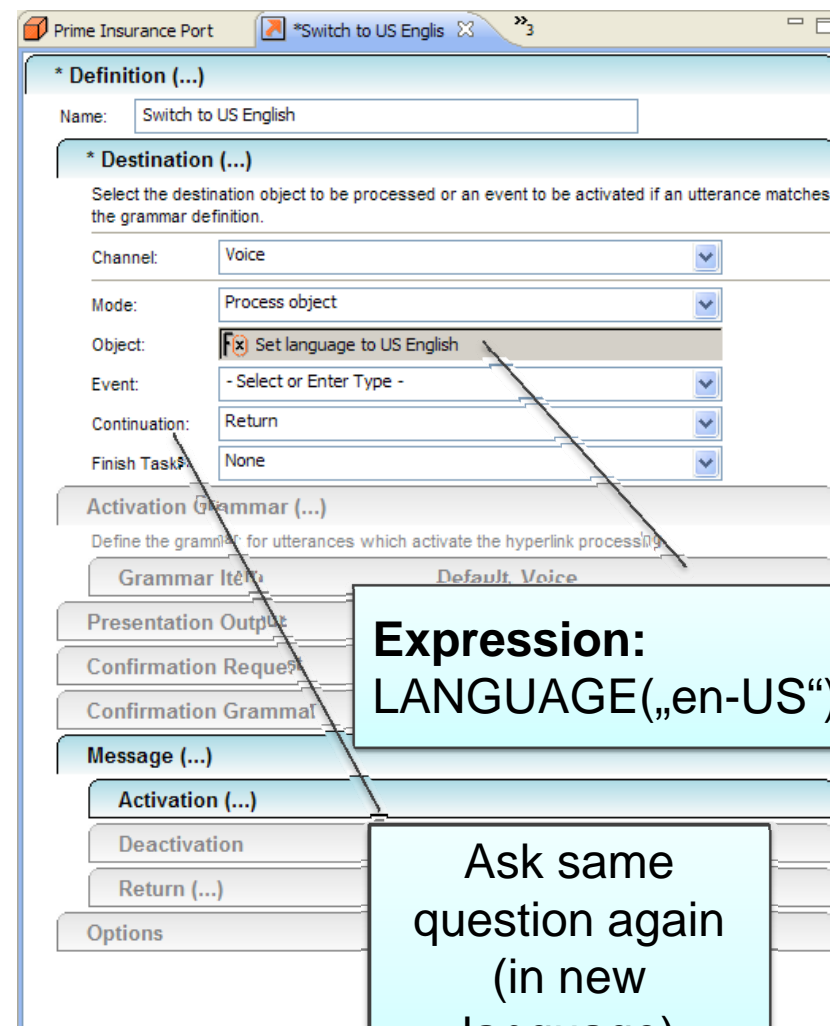
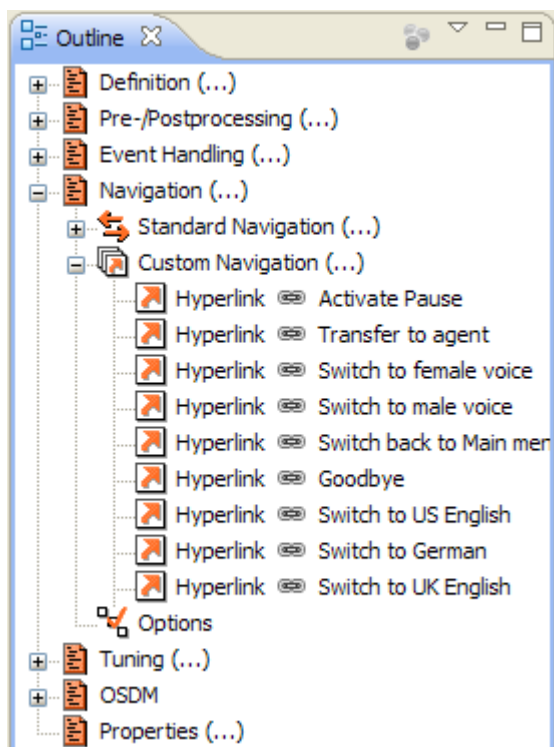
```
<?xml version="1.0" encoding="ISO-8859-1"?>
<grammar version="1.0"
  xmlns="http://www.w3.org/2001/06/grammar"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.w3.org/2001/06/grammar http://www.w3.org/TR/speech-grammar/grammar.xsd"
  xml:lang="en-US" mode="voice" tag-format="semantics/1.0" root="Language">
  <rule id="Language" scope="public">
    <one-of>
      <item>
        <one-of>
          <item xml:lang="en-US">english</item>
          <item xml:lang="es-MX">inglés</item>
        </one-of>
        <tag sltLanguage="en-US" </tag>
      </item>
      <item>
        <one-of>
          <item xml:lang="en-US">spanish</item>
          <item xml:lang="es-MX">español</item>
        </one-of>
        <tag sltLanguage="es-MX" </tag>
      </item>
    </one-of>
  </rule>
</grammar>
```

Mark different Languages with the *xml:lang* attribute



In-session language switch

- Configure „language-switch“ *Hyperlinks* in the root *Module*





Agenda

- Managing multilingual prompts
 - Audio resource management best practices
 - Adding new languages with Storyboard Manager
 - Accounting for cultural and syntactic differences
- Let the caller choose
 - Initial / in-session language selection
- Natural pronunciation of dynamic data
 - The Formatting Bus (aka TTA Bus)

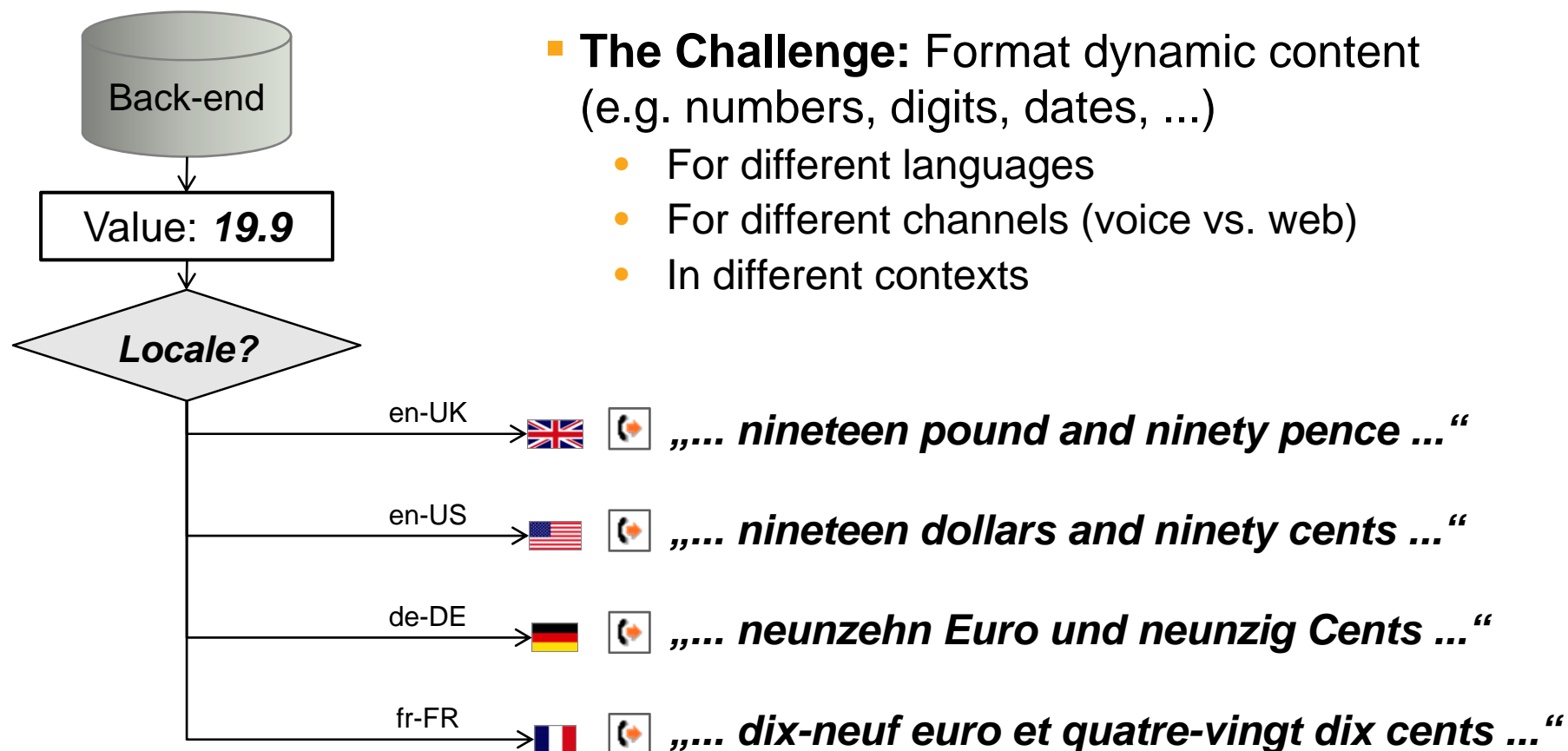


Analysis and Reporting

- Deployment Scenarios
 - Multi-language / multi-country deployments



Natural Pronunciation

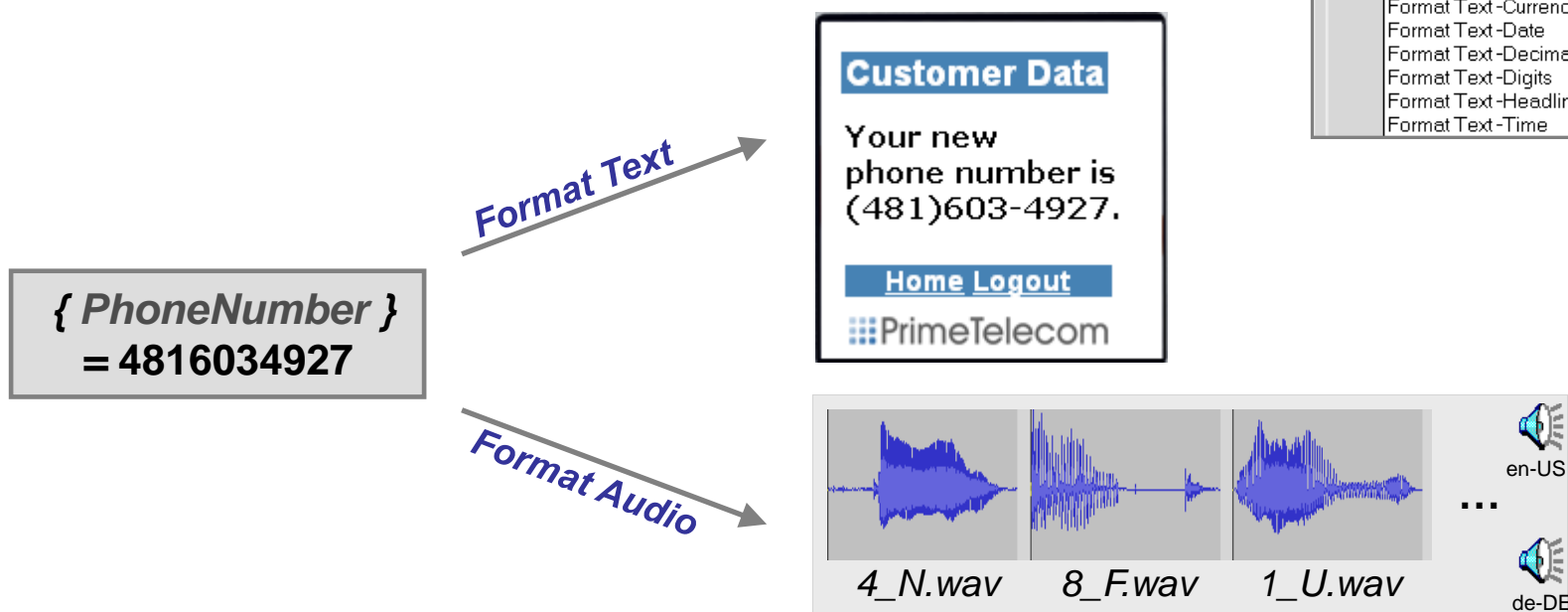
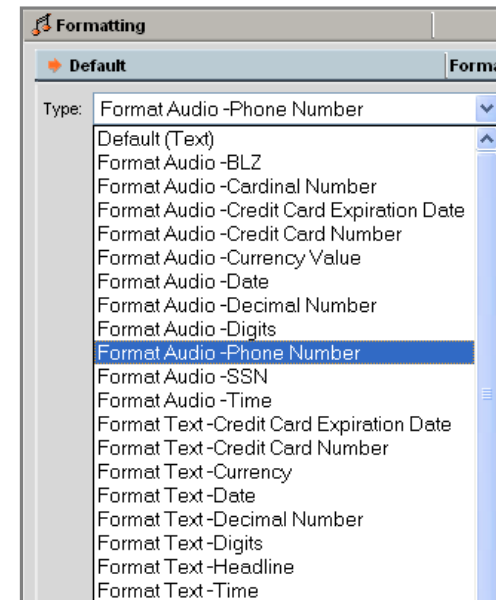


- **The Challenge:** Format dynamic content (e.g. numbers, digits, dates, ...)
 - For different languages
 - For different channels (voice vs. web)
 - In different contexts

Natural Dialog Management

Natural Pronunciation

- The VoiceObjects “Formatting Bus”
 - Natural pronunciation of dynamic content
 - Best-practice standard formats provided
 - *Date, Time, Digits, Number, Currency, Credit Card, ...*
 - Custom formatting can be added

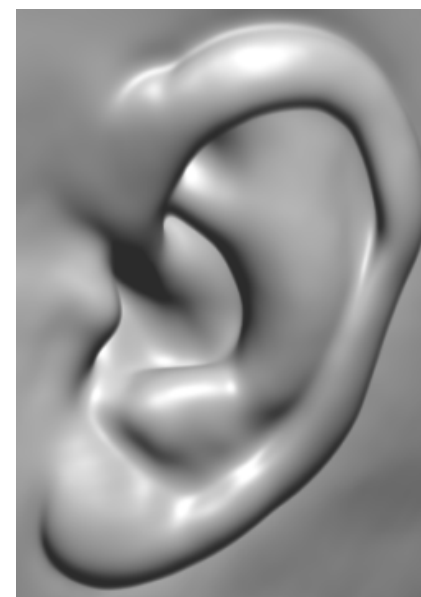




Formatting Bus - Specification

- **Step 1:** Define sample data
 - Typical exceptions / variations, e.g. *hundred – a hundred – one hundred, hundred five – hundred and five*
 - Different groupings of digits, e.g. *1-2-3 4-5-6*
>> monotone / melodic prosody
 - Data in different contexts, e.g. *un pesos* vs. *cinco – zero – uno*

- **Step 2:** Listen carefully to native speakers & discuss with native linguists
 - Intonation, stress, pitch, pauses etc.
 - Usage (register, style etc.)





Formatting Bus - Specification

- **Step 3:** Extract concatenation rules

■ **threehundred_O million_O Euros_D** (... 300.000.000 €.)

■ **ten_F million_U fivehundred_O Euros_D** (... 10.000.500 €.)

- **Step 4:** Implementation of concatenation rules

- **Step 5:** Test algorithms against sample data

8_F.wav 

8_O.wav 


8_U.wav 


8_D.wav 



Formatting Bus - Examples

1,254,000

English 


Spanish 

Netherlands 

Finnish 


Dates in French

September 28 

October 28 

Mid- / End-Sentence

4 £ 

4 £ 50 p. 



Formatting Bus – XML-TTA Format

Data: 123.01

Processing as: Currency

Output:

Language = en-US
Channel = **voice**

Language = de-DE
Channel = **web**

```

<root>
  <row>
    <col name="file">100_O</col>
  </row>
  <row>
    <col name="file">23_O</col>
    <col name="text"><![CDATA[123]]></col>
  </row>
  <row>
    <col name="file">dollarsAnd</col>
    <col name="text"><![CDATA[dollars and]]></col>
  </row>
  <row>
    <col name="file">1_O</col>
    <col name="text"><![CDATA[1]]></col>
  </row>
  <row>
    <col name="file">cent_D</col>
    <col name="text"><![CDATA[cent]]></col>
  </row>
</root>

```



```

<root>
  <row>
    <col name="markup"><![CDATA[123.01 €]]>
  </row>
</root>

```





Formatting Bus Configuration

The image shows a configuration process in SciTE. The main window displays the XML code for `Format-TTA.xml`. The `<srcPath>` attribute is set to `./lib/ttaLibrary/`. A file explorer window shows the directory structure, with the `lib` folder and `ttaLibrary` subfolder highlighted. A configuration dialog for 'Item Count' is open, showing the 'Format Item' dropdown set to 'Format Audio -Cardinal Number, Default'.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<ttaLibrary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation="../schemas/TTALib"
  <header>
    <defaultLocaleID>en</defaultLocaleID>
    <namePrefix><locale id="en">Format Audio - </locale>
    <srcPath>./lib/ttaLibrary/</srcPath>
  </header>
  <function>
```

File Explorer: `C:\VoiceObjectsDesktop\plugins\com.voiceobjects.eclipseDesktop_7.3.0\lib\ttaLibrary`

Configuration Dialog: `Item Count`

- Name: Item Count
- Declaration: [Empty]
- Formatting (...): Define the formatting options for playback.
- Format Item: Format Audio -Cardinal Number, Default
- Label: [Empty]
- Layer: [Empty]
- Channel: Default
- Language: Default
- Type: Format Audio -Cardinal Number
- Options: end=true
- Audio location: Audio Locator (TTA)
- File prefix: [Empty]
- File suffix: [Empty]
- File extension: .wav
- Random variations: Disabled
- Intermediate silence: Disabled
- Value substitution: [Empty]
- Custom separators: [Empty]



Formatting Bus Configuration

```
Format-TTA.xml * SciTE
File Edit Search View Tools Options Language Buffers Help
Format-TTA.xml *
<?xml version="1.0" encoding="ISO-8859-1"?>
<ttaLibrary xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation="../schemas/TTALibrary.xsd">
  <header>
  <function>
    <properties>
    <java>
      <srcFile>ttalib.jar</srcFile>
      <class>com.voiceobjects.TTABus.TTACardinalNumber</class>
      <method>process</method>
      <parameterSet>
        <result>XMLResult</result>
        <data>data</data>
        <resourceLocator>resourceLocator</resourceLocator>
        <language>language</language>
        <encoding>encoding</encoding>
        <options>options</options>
        <channel>channel</channel>
        <pronunciation>pronunciation</pronunciation>
      </parameterSet>
    </java>
    <filePostProcessing>
  </function>
```

- TTACardinalNumber
 - TTACardinalNumber()
 - validateOptions()
 - process()
 - setChannel(String)
 - setData(String)
 - setPronunciation(String)
 - setResourceLocator(String)
 - setLanguage(String)
 - setEncoding(String)
 - setOptions(String)
 - setCustomSeparator(String)
 - setErrorMsg(String)
 - getXMLResult()



Agenda

- Managing multilingual prompts
 - Audio resource management best practices
 - Adding new languages with Storyboard Manager
 - Accounting for cultural and syntactic differences
- Let the caller choose
 - Initial / in-session language selection
- Natural pronunciation of dynamic data
 - The Formatting Bus (aka TTA Bus)



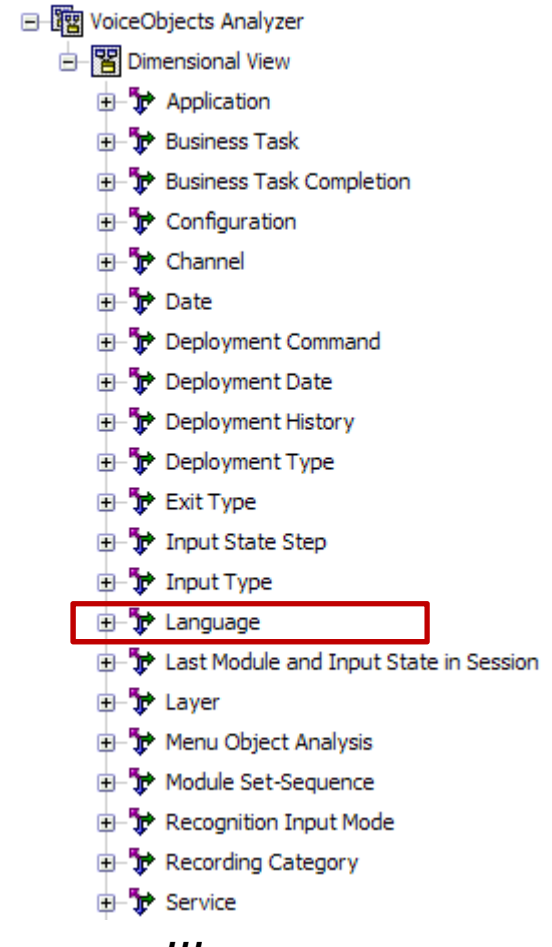
Analysis and Reporting

- Deployment Scenarios
 - Multi-language / multi-country deployments



Analysis and Reporting

- **VoiceObjects automatically logs the dialog language for every session.**
- Every report in VO Analyzer can be configured to use the dialog language
 - ... for grouping results
 - ... as a query filter: Language
- Analyze, by dialog language:
 - Speech Recognition success
 - Navigation patterns
 - Business task completion rates
 - Layer Statistics
 - ...





Agenda

- Managing multilingual prompts
 - Audio resource management best practices
 - Adding new languages with Storyboard Manager
 - Accounting for cultural and syntactic differences
- Let the caller choose
 - Initial / in-session language selection
- Natural pronunciation of dynamic data
 - The Formatting Bus (aka TTA Bus)



Analysis and Reporting

- Deployment Scenarios
 - Multi-language / multi-country deployments

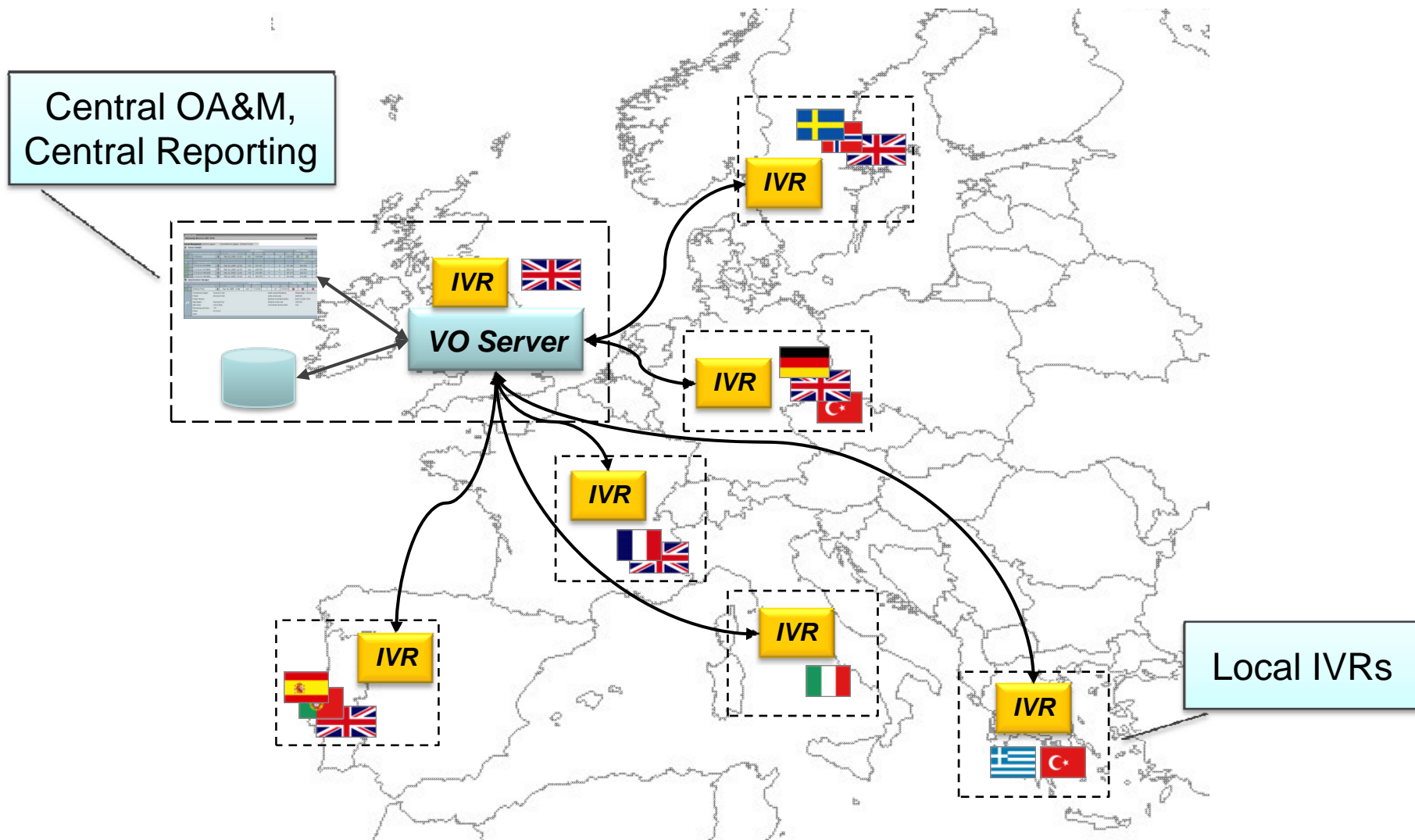


Multi-country / Multi-language





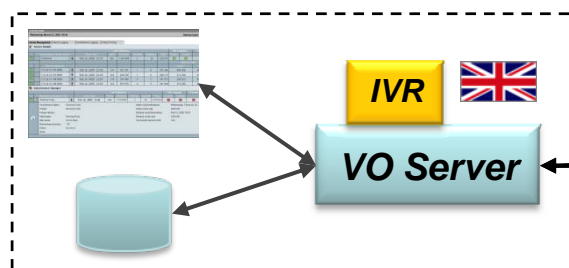
Central Control



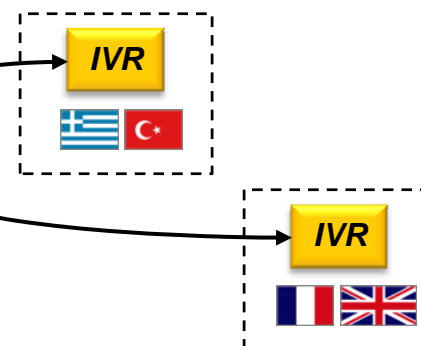


Central Control

■ Central Site:



■ Local Sites:



● Software

- VO Server Metadata Repository and Logging DB (multi-tenant)
- BI Infrastructure

● Tasks

- Full administration of central and local sites (local access to VO Desktop)
- Reporting on full system performance (across local sites)

● Software

- IVR (with VoiceXML Browser)
- Web Server with Audio and Grammar resources

● Tasks

- Administration of local site (remote access to VO Desktop; multi-tenant access control)
- Reporting on local system performance (via remote access to BI solution)



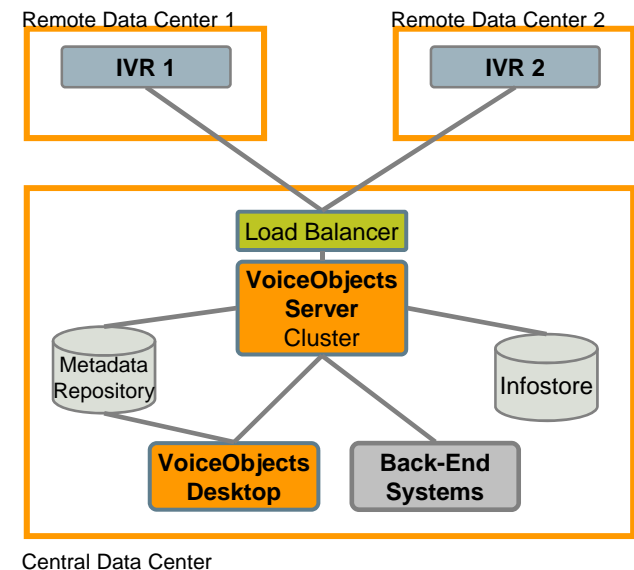
Central Deployment - Benefits

■ Benefits

- **Central management** and operations
- **Resilience** against failures of VO Server Instances
- **Consolidated logging** to central Infostore DB (no merging of log data necessary)
- **Low bandwidth requirements** due to small VXML pages (~2-5 kB) rendered by VO Server
- **Virtualization** for remote data centers through Multi-tenancy

■ Caveat

- Possible network *latency* issues in VXML traffic between IVR and VO Server (network latency SLAs required)





Thank You!

Questions

