



# Intorduction to VoiceXML (3)

# Main Topics

- Delivering Content
- Collecting User Input

# <prompt>

```
<vxml version="2.1" xmlns="http://www.w3.org/2001/vxml">  
  <form>  
    <block>  
      <prompt> Hello, world!  
    </prompt>  
  </block>  
</form>  
</vxml>
```

- The <prompt> element has several attributes
  - bargein – controls barge-in, for user to interrupt the prompt before it is complete
  - count: allows the specification of different prompts if the dialog is going to be executed repeatedly
  - timeout: The number of seconds (s) or milliseconds (ms) the platform waits for user input before throwing a noinput event.

# <audio>

- Pre-recorded audio is played inside a prompt with the <audio> element.
- The src attribute is required to specify the URI of the audio file to be played.
- The <audio> element may include an alternative TTS message between the opening and closing tags,
  - ☞ to be played in case the retrieval of the audio file fails, or the file itself is corrupted

<prompt>

Hello! Thanks for calling

<audio src="ncnu.wav">National Chi Nan University</audio>.

<audio src="welcome.wav"/>

</prompt>

# Improving TTS with Speech Markup

- <break>
- <emphasis>
- <say-as>
- <paragraph>, <sentence>
- <prosody>

# <break>

- Introduces a specified amount of silence into the audio played back to the user.
- If the units are not specified, the VoiceXML interpreter assumes milliseconds.

<block>

<prompt>

Good morning!

<break time="1.5s"/>

Welcome to

<break time="500"/>

News of the Day

<break size="medium"/>

</prompt>

<exit/>

</block>

# Size of <break>

Value of attribute “size”	Pause
none	0ms
small	200ms
medium	500ms
large	1000ms

# <emphasis>

- Indicates the prominence or stress with which the enclosed text should be played back by the TTS engine.

```
<prompt bargein="true">
```

```
You can say <emphasis level="strong">help</emphasis>  
at any time.
```

```
<break size="small"/>
```

```
You can say <emphasis level="strong">operator</emphasis>  
to be transferred to a representative between 8 AM and 5PM.
```

```
<break size="5500ms"/>
```

```
Say <emphasis>status</emphasis> to check your account.
```

```
</prompt>
```

- Currently the **emphasis** element has **no effect** on the synthesized speech output.



## <say-as>

- Aids the TTS engine in pronunciation by resolving ambiguities regarding the meaning of the contained text.

<prompt>

<say-as type="address"> 1310 Villa St </say-as>

<say-as type="address"> 20 Greenwood Rd </say-as>

<say-as type="time"> 9:00 AM </say-as>

<say-as type="time"> 5:00 PM </say-as>

</prompt>

# <sentence>

- Identifies the enclosed text as a complete sentence.

<prompt>

<sentence>This is sentence one.</sentence>

<sentence>This is sentence two.</sentence>

</prompt>

# <paragraph>

- Provides a container element for zero or more sentences.

<prompt>

<paragraph>This is paragraph one.</paragraph>

<paragraph>This is paragraph two.</paragraph>

</prompt>

# <prosody>

- Controls the speaking rate and the volume of the speech output.
- `<prosody rate="string" volume="string"/>`
- **rate**
  - fast - Applies a fast speaking rate when playing back the contained text.
  - medium - Applies a medium speaking rate when playing back the contained text.
  - slow - Applies a slow speaking rate when playing back the contained text.
  - default - Applies the default speaking rate when playing back the contained text.
- **volume**
  - silent - Mutes output when playing back the contained text.
  - soft - Lowers the volume when playing back the contained text.
  - medium - Applies medium volume when playing back the contained text.
  - loud - Raises the volume when playing back the contained text.
  - default - Uses the default volume when playing back the contained text.

# Prosody.vxml

```
<vxml version="2.0">
  <form>
    <block>
      <prompt>
        <prosody rate="slow" volume="loud">
          This is a slow, loud prompt.
        </prosody>
      </prompt>
    <exit/>
  </block>
</form>
</vxml>
```

# Collecting User Input

- <field>
- <record>

# <field>

- Formulates an interactive dialog between the user and the system.

```
<vxml version="2.0">
  <form>
    <field name="ph" type="digits">
      <prompt>
        What is your phone number?
      </prompt>
      <filled>
        <prompt> You said <value expr="ph"/>.</prompt>
      </filled>
    </field>
  </form>
</vxml>
```

# Built-in Grammar

```
<vxml version="2.0">
  <form>
    <field name="answer" type="boolean">
      <nomatch>
        Please say YES or NO.
      </nomatch>
      <noinput>
        Sorry. I did not hear you. Please try to speak more loudly.
      </noinput>
      <prompt>
        Are you going to school today?
      </prompt>
      <filled>
        <prompt> You said <value expr="answer"/>.</prompt>
      </filled>
    </field>
  </form>
</vxml>
```



# Field-2.vxml

- Computer: Please tell me a number.
- User: ...
- Computer: Sorry. I did not hear you. Please tell me a number.
- User: 7
- Computer: You said “seven”. The square of seven is forty-nine.

# <record>

## ■ Records user input.

```
<vxml version="2.0">
  <form>
    <record name="msg" beep="true" finalsilence="2000ms"
      dtmfterm="true">
      <prompt>
        Please leave your message after the beep.
      </prompt>
    </record>
    <block>
      <prompt> You said <value expr="msg"/>.</prompt>
    </block>
  </form>
</vxml>
```

# Ncnu.vxml

- 國立暨南國際大學您好. 資訊系請按一, 通訊所請按二, 電機系請按三...