

# Text-to-Speech Synthesis of Irish

## Project Overview

Amelia Kelly

Phonetics & Speech Laboratory  
Centre for Language and Communication Studies  
University of Dublin, Trinity College

February 2009 / University of Amsterdam

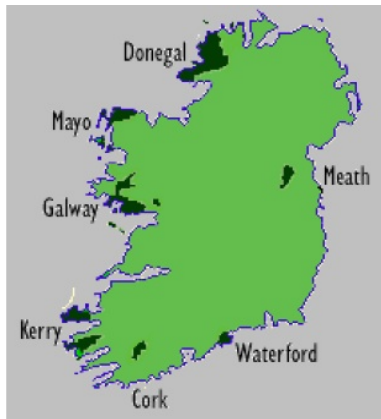
# Outline

- 1 The Irish Language
  - Overview of the Irish Language
  - Applications of Minority Language Synthesis
- 2 Methods of Concatenative Speech Synthesis
  - Pre-Requisites for Corpus-Based Synthesis
  - Unit Selection Synthesis
  - HMM Synthesis
- 3 The `abair.ie` Project
  - Unit Selection Synthesis of Irish Dialects

# Outline

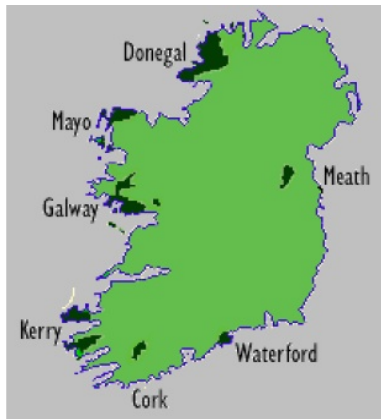
- 1 The Irish Language
  - Overview of the Irish Language
  - Applications of Minority Language Synthesis
- 2 Methods of Concatenative Speech Synthesis
  - Pre-Requisites for Corpus-Based Synthesis
  - Unit Selection Synthesis
  - HMM Synthesis
- 3 The `abair.ie` Project
  - Unit Selection Synthesis of Irish Dialects

# The Irish Language in Ireland



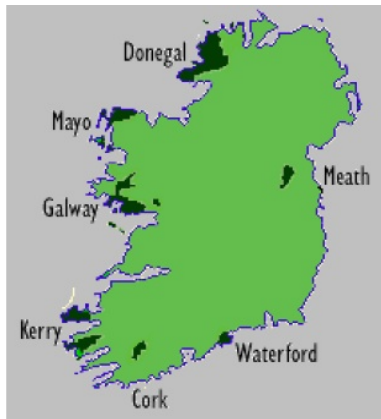
- One of the oldest European Languages
- Ireland's first language and an official language of the EU
- Spoken natively in only a few areas of the country
- Compulsory school subject

# The Irish Language in Ireland



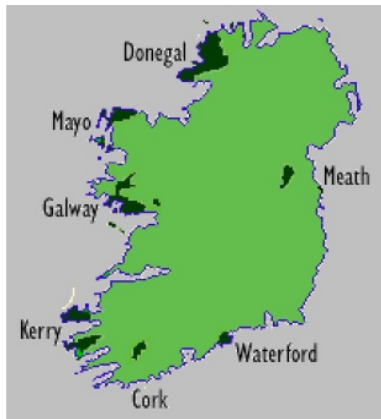
- One of the oldest European Languages
- Ireland's first language and an official language of the EU
- Spoken natively in only a few areas of the country
- Compulsory school subject

# The Irish Language in Ireland



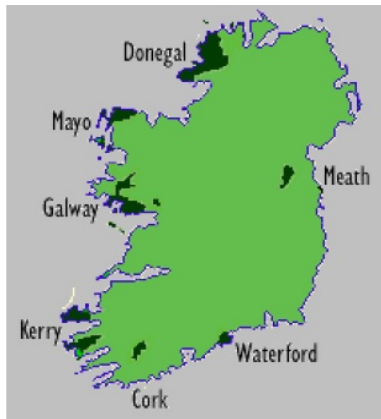
- One of the oldest European Languages
- Ireland's first language and an official language of the EU
- Spoken natively in only a few areas of the country
- Compulsory school subject

# The Irish Language in Ireland



- One of the oldest European Languages
- Ireland's first language and an official language of the EU
- Spoken natively in only a few areas of the country
- Compulsory school subject

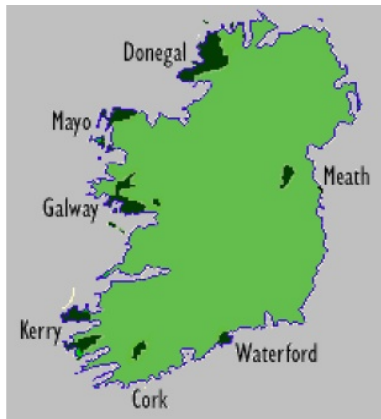
# The Irish Language in Ireland



- One of the oldest European Languages
- Ireland's first language and an official language of the EU
- Spoken natively in only a few areas of the country
- Compulsory school subject

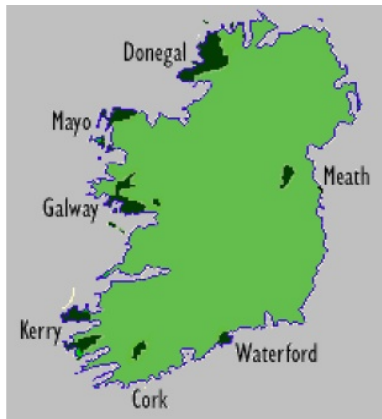


# The Irish Language in Ireland



- One of the oldest European Languages
- Ireland's first language and an official language of the EU
- Spoken natively in only a few areas of the country
- Compulsory school subject

# The Irish Language in Ireland



- One of the oldest European Languages
- Ireland's first language and an official language of the EU
- Spoken natively in only a few areas of the country
- Compulsory school subject

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaigh* is pronounced [nʲiː wiː]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [báːd] means *boat*; *báid* [báːdʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaigh* is pronounced [nʲiː wʲiː]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [bád] means *boat*; *báid* [bádʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaigh* is pronounced [nʲiː wʲiː]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [bád] means *boat*; *báid* [bádʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaigh* is pronounced [nʲiː wʲiː]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [báːd] means *boat*; *báid* [báːdʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaigh* is pronounced [nʲiː wʲiː]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [báːd] means *boat*; *báid* [báːdʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaigh* is pronounced [nʲiː wʲiː]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [báːd] means *boat*; *báid* [báːdʲ] means *boats*



# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaighigh* is pronounced [nʲi: wi:]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [bád] means *boat*; *báid* [báidʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaighigh* is pronounced [nʲi: wi:]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [bád] means *boat*; *báid* [báidʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaighigh* is pronounced [nʲi: wi:]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [bád] means *boat*; *báid* [bádʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaighigh* is pronounced [nʲi: wi:]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [bád] means *boat*; *báid* [bádʲ] means *boats*

# The Difficulty with Irish

- 1 Lack of interaction with native speakers
  - Many teachers are not native or fluent speakers
  - Very few remaining Gaeltacht regions
- 2 Lack of a standard dialect
  - Dialect differs drastically from place to place
  - Difficult to understand a speaker of a different dialect
- 3 Complex orthographical system
  - *Ní bhfaighigh* is pronounced [n<sup>j</sup>i: wi:]
- 4 Complex sound system
  - palatalised and velarised consonant oppositions
  - *bád* [bád] means *boat*; *báid* [bád<sup>j</sup>] means *boats*

# Irish Consonants

	Labial	Dental	Alveolar	Alveolo-palatal	Palatal	Velar	Glottal
Plosive	p <sup>ʲ</sup> b <sup>ʲ</sup> p <sup>ˠ</sup> b <sup>ˠ</sup>	t <sup>ʲ</sup> d <sup>ʲ</sup> t <sup>ˠ</sup> d <sup>ˠ</sup>		t <sup>ʲ</sup> d <sup>ʲ</sup> t <sup>ˠ</sup> d <sup>ˠ</sup>	c j ç ʝ	k g	
Fricative/ Approximant	f <sup>ʲ</sup> w f <sup>ˠ</sup> v <sup>ˠ</sup>		s <sup>ʲ</sup>	ç	ç j x ɣ	x ɣ	h
Nasal	m <sup>ʲ</sup> m <sup>ˠ</sup>	n <sup>ʲ</sup>	n	n <sup>ʲ</sup>	ɲ	ŋ	
Tap			r <sup>ʲ</sup> r <sup>ˠ</sup>				
Lateral Approximant		l <sup>ʲ</sup>	l	l <sup>ʲ</sup>			

# Outline

- 1 The Irish Language
  - Overview of the Irish Language
  - Applications of Minority Language Synthesis
- 2 Methods of Concatenative Speech Synthesis
  - Pre-Requisites for Corpus-Based Synthesis
  - Unit Selection Synthesis
  - HMM Synthesis
- 3 The `abair.ie` Project
  - Unit Selection Synthesis of Irish Dialects

# Some Applications

## Language Learning

- Teaching aids and language learning devices

## Accessibility

- Screen readers for the visually impaired
- Synthetic voices for the vocally impaired

## Other Applications

- Synthetic voices for online gaming avatars
- Historical archives - each voice is a snapshot of an Irish dialect at a particular time



# Some Applications

## Language Learning

- Teaching aids and language learning devices

## Accessibility

- Screen readers for the visually impaired
- Synthetic voices for the vocally impaired

## Other Applications

- Synthetic voices for online gaming avatars
- Historical archives - each voice is a snapshot of an Irish dialect at a particular time

# Some Applications

## Language Learning

- Teaching aids and language learning devices

## Accessibility

- Screen readers for the visually impaired
- Synthetic voices for the vocally impaired

## Other Applications

- Synthetic voices for online gaming avatars
- Historical archives - each voice is a snapshot of an Irish dialect at a particular time

# Some Applications

## Language Learning

- Teaching aids and language learning devices

## Accessibility

- Screen readers for the visually impaired
- Synthetic voices for the vocally impaired

## Other Applications

- Synthetic voices for online gaming avatars
- Historical archives - each voice is a snapshot of an Irish dialect at a particular time

# Some Applications

## Language Learning

- Teaching aids and language learning devices

## Accessibility

- Screen readers for the visually impaired
- Synthetic voices for the vocally impaired

## Other Applications

- Synthetic voices for online gaming avatars
- Historical archives - each voice is a snapshot of an Irish dialect at a particular time

# Some Applications

## Language Learning

- Teaching aids and language learning devices

## Accessibility

- Screen readers for the visually impaired
- Synthetic voices for the vocally impaired

## Other Applications

- Synthetic voices for online gaming avatars
- Historical archives - each voice is a snapshot of an Irish dialect at a particular time

# Some Applications

## Language Learning

- Teaching aids and language learning devices

## Accessibility

- Screen readers for the visually impaired
- Synthetic voices for the vocally impaired

## Other Applications

- Synthetic voices for online gaming avatars
- Historical archives - each voice is a snapshot of an Irish dialect at a particular time

# Some Applications

## Language Learning

- Teaching aids and language learning devices

## Accessibility

- Screen readers for the visually impaired
- Synthetic voices for the vocally impaired

## Other Applications

- Synthetic voices for online gaming avatars
- Historical archives - each voice is a snapshot of an Irish dialect at a particular time

# Outline

- 1 The Irish Language
  - Overview of the Irish Language
  - Applications of Minority Language Synthesis
- 2 **Methods of Concatenative Speech Synthesis**
  - **Pre-Requisites for Corpus-Based Synthesis**
  - Unit Selection Synthesis
  - HMM Synthesis
- 3 The `abair.ie` Project
  - Unit Selection Synthesis of Irish Dialects



# The Corpus

- Large database of recorded speech, labelled as units
- Many examples of each speech unit in many different contexts
- The bigger the corpus, the higher the quality of the synthesised speech

## Example of a Uniphone Corpus

- "A whole joy was reaping."
- "But they've gone south."
- "You should fetch azure Mike."

# The Corpus

- Large database of recorded speech, labelled as units
- Many examples of each speech unit in many different contexts
- The bigger the corpus, the higher the quality of the synthesised speech

## Example of a Uniphone Corpus

- "A whole joy was reaping."
- "But they've gone south."
- "You should fetch azure Mike."

# The Corpus

- Large database of recorded speech, labelled as units
- Many examples of each speech unit in many different contexts
- The bigger the corpus, the higher the quality of the synthesised speech

## Example of a Uniphone Corpus

- "A whole joy was reaping."
- "But they've gone south."
- "You should fetch azure Mike."

# The Corpus

- Large database of recorded speech, labelled as units
- Many examples of each speech unit in many different contexts
- The bigger the corpus, the higher the quality of the synthesised speech

## Example of a Uniphone Corpus

- "A whole joy was reaping."
- "But they've gone south."
- "You should fetch azure Mike."

# The Corpus

- Large database of recorded speech, labelled as units
- Many examples of each speech unit in many different contexts
- The bigger the corpus, the higher the quality of the synthesised speech

## Example of a Uniphone Corpus

- "A whole joy was reaping."
- "But they've gone south."
- "You should fetch azure Mike."

# The Corpus

- Large database of recorded speech, labelled as units
- Many examples of each speech unit in many different contexts
- The bigger the corpus, the higher the quality of the synthesised speech

## Example of a Uniphone Corpus

- "A whole joy was reaping."
- "But they've gone south."
- "You should fetch azure Mike."

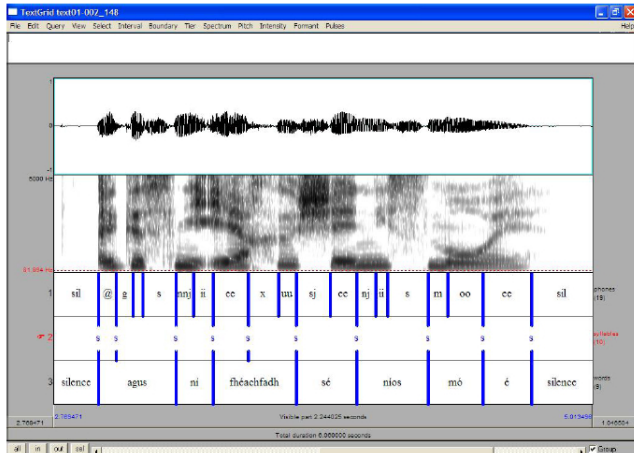
# The Corpus

- Large database of recorded speech, labelled as units
- Many examples of each speech unit in many different contexts
- The bigger the corpus, the higher the quality of the synthesised speech

## Example of a Uniphone Corpus

- "A whole joy was reaping."
- "But they've gone south."
- "You should fetch azure Mike."

# Annotation & Alignment

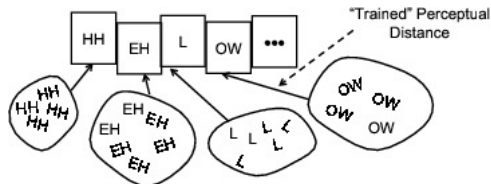




# Outline

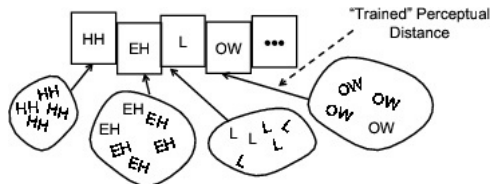
- 1 The Irish Language
  - Overview of the Irish Language
  - Applications of Minority Language Synthesis
- 2 **Methods of Concatenative Speech Synthesis**
  - Pre-Requisites for Corpus-Based Synthesis
  - **Unit Selection Synthesis**
  - HMM Synthesis
- 3 The `abair.ie` Project
  - Unit Selection Synthesis of Irish Dialects

# How Unit Selection Synthesis Works



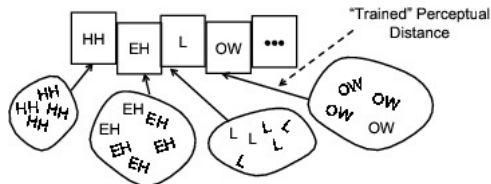
- The **target utterance** is the sequence of units to be synthesised
- A set of candidates are chosen for each unit in the target utterance
- The most appropriate sequence of units is chosen by calculating the **cost function**

# How Unit Selection Synthesis Works



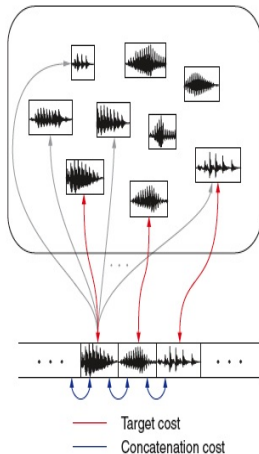
- The **target utterance** is the sequence of units to be synthesised
- A set of candidates are chosen for each unit in the target utterance
- The most appropriate sequence of units is chosen by calculating the **cost function**

# How Unit Selection Synthesis Works



- The **target utterance** is the sequence of units to be synthesised
- A set of candidates are chosen for each unit in the target utterance
- The most appropriate sequence of units is chosen by calculating the **cost function**

# Target and Concatenation Cost



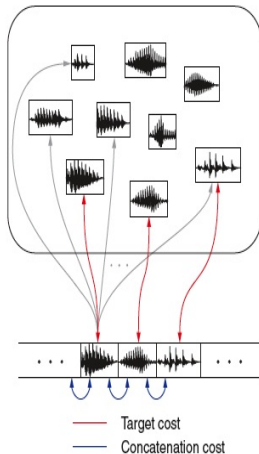
## ● Target Cost

How similar is the unit to the target?

## ● Concatenation Cost

How similar are the units to each other?

# Target and Concatenation Cost



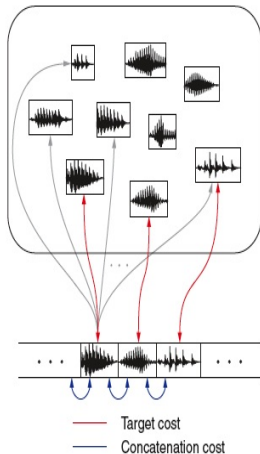
## ● Target Cost

How similar is the unit to the target?

## ● Concatenation Cost

How similar are the units to each other?

# Target and Concatenation Cost



- Target Cost

- Similarity between the target unit and the candidate unit

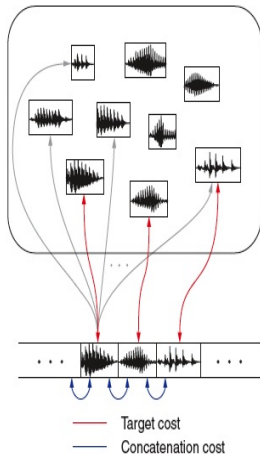
$$C^t(t_i, u_i) = \sum_{j=1}^p w_j^t C_j^t(t_i, u_i)$$

- Concatenation Cost

- Perceptual similarity between one chosen unit and the next

$$C^c(u_{i-1}, u_i) = \sum_{j=1}^q w_c^k C_c^k(u_{i-1}, u_i)$$

# Target and Concatenation Cost



## • Target Cost

- Similarity between the target unit and the candidate unit

$$C^t(t_i, u_i) = \sum_{j=1}^p w_j^t C_j^t(t_i, u_i)$$

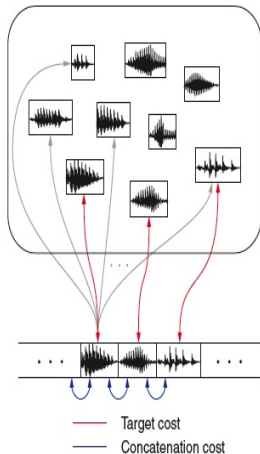
## • Concatenation Cost

- Perceptual similarity between one chosen unit and the next

$$C^c(u_{i-1}, u_i) = \sum_{j=1}^q w_c^k C_c^k(u_{i-1}, u_i)$$



# Target and Concatenation Cost



- Target Cost

- Similarity between the target unit and the candidate unit

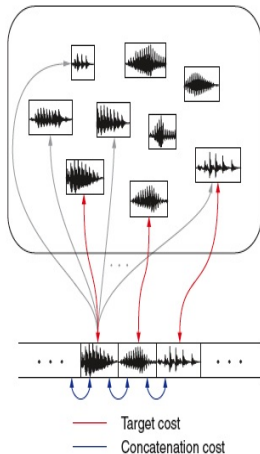
$$C^t(t_i, u_i) = \sum_{j=1}^p w_j^t C_j^t(t_i, u_i)$$

- Concatenation Cost

- Perceptual similarity between one chosen unit and the next

$$C^c(u_{i-1}, u_i) = \sum_{j=1}^q w_j^c C_j^c(u_{i-1}, u_i)$$

# Target and Concatenation Cost



- Target Cost

- Similarity between the target unit and the candidate unit

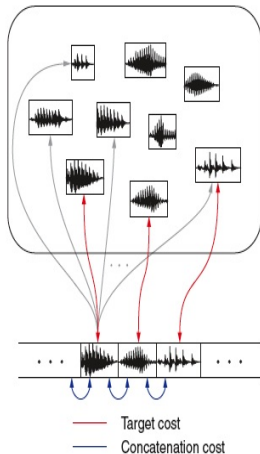
$$C^t(t_i, u_i) = \sum_{j=1}^p w_j^t C_j^t(t_i, u_i)$$

- Concatenation Cost

- Perceptual similarity between one chosen unit and the next

$$C^c(u_{i-1}, u_i) = \sum_{j=1}^q w_c^k C_c^k(u_{i-1}, u_i)$$

## Target and Concatenation Cost



- Target Cost

- Similarity between the target unit and the candidate unit

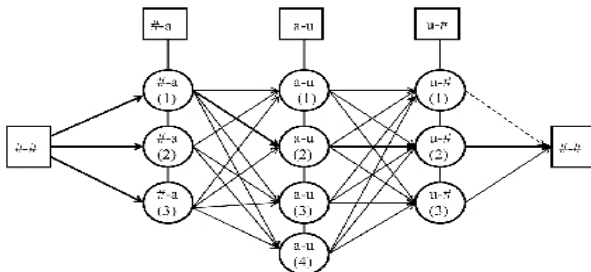
$$C^t(t_i, u_i) = \sum_{j=1}^p w_j^t C_j^t(t_i, u_i)$$

- Concatenation Cost

- Perceptual similarity between one chosen unit and the next

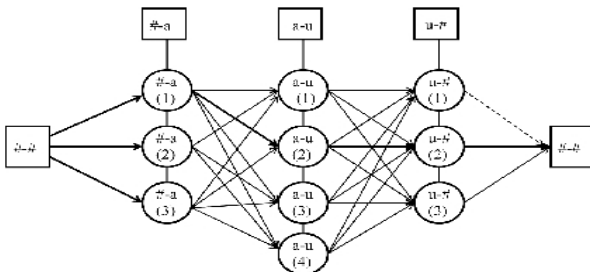
$$C^c(u_{i-1}, u_i) = \sum_{j=1}^q w_c^k C_c^k(u_{i-1}, u_i)$$

# Viterbi Search



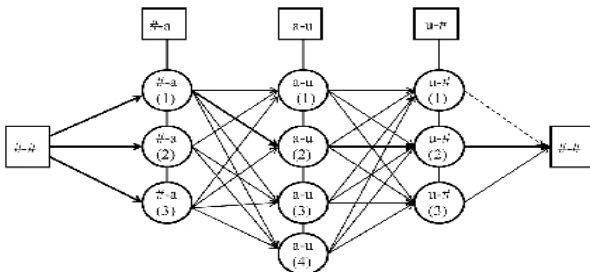
- Once the costs have been calculated, the best sequence is selected using a Viterbi search
- Calculates a cost for every possible combination of units
- The units selected are those which combine to give the sentence with the lowest cost

# Viterbi Search



- Once the costs have been calculated, the best sequence is selected using a Viterbi search
- Calculates a cost for every possible combination of units
- The units selected are those which combine to give the sentence with the lowest cost

# Viterbi Search



- Once the costs have been calculated, the best sequence is selected using a Viterbi search
- Calculates a cost for every possible combination of units
- The units selected are those which combine to give the sentence with the lowest cost

# Pros and cons of Unit Selection Synthesis

## Pros

- High quality, natural sounding speech
- Prosody and intonation are preserved
- Minimum signal processing required

## Cons

- Large database needed containing hours of recorded speech
- Synthesis can be very slow
- Voice cannot be changed without sacrificing quality

# Pros and cons of Unit Selection Synthesis

## Pros

- High quality, natural sounding speech
- Prosody and intonation are preserved
- Minimum signal processing required

## Cons

- Large database needed containing hours of recorded speech
- Synthesis can be very slow
- Voice cannot be changed without sacrificing quality



# Pros and cons of Unit Selection Synthesis

## Pros

- High quality, natural sounding speech
- Prosody and intonation are preserved
- Minimum signal processing required

## Cons

- Large database needed containing hours of recorded speech
- Synthesis can be very slow
- Voice cannot be changed without sacrificing quality

# Pros and cons of Unit Selection Synthesis

## Pros

- High quality, natural sounding speech
- Prosody and intonation are preserved
- Minimum signal processing required

## Cons

- Large database needed containing hours of recorded speech
- Synthesis can be very slow
- Voice cannot be changed without sacrificing quality

# Pros and cons of Unit Selection Synthesis

## Pros

- High quality, natural sounding speech
- Prosody and intonation are preserved
- Minimum signal processing required

## Cons

- Large database needed containing hours of recorded speech
- Synthesis can be very slow
- Voice cannot be changed without sacrificing quality

# Pros and cons of Unit Selection Synthesis

## Pros

- High quality, natural sounding speech
- Prosody and intonation are preserved
- Minimum signal processing required

## Cons

- Large database needed containing hours of recorded speech
- Synthesis can be very slow
- Voice cannot be changed without sacrificing quality

# Pros and cons of Unit Selection Synthesis

## Pros

- High quality, natural sounding speech
- Prosody and intonation are preserved
- Minimum signal processing required

## Cons

- Large database needed containing hours of recorded speech
- Synthesis can be very slow
- Voice cannot be changed without sacrificing quality

# Pros and cons of Unit Selection Synthesis

## Pros

- High quality, natural sounding speech
- Prosody and intonation are preserved
- Minimum signal processing required

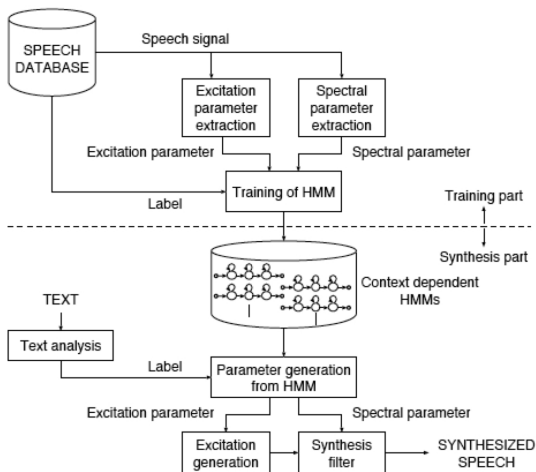
## Cons

- Large database needed containing hours of recorded speech
- Synthesis can be very slow
- Voice cannot be changed without sacrificing quality

# Outline

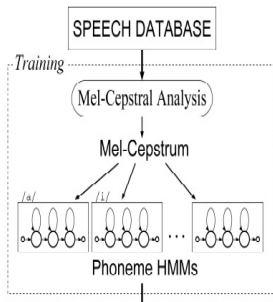
- 1 The Irish Language
  - Overview of the Irish Language
  - Applications of Minority Language Synthesis
- 2 **Methods of Concatenative Speech Synthesis**
  - Pre-Requisites for Corpus-Based Synthesis
  - Unit Selection Synthesis
  - **HMM Synthesis**
- 3 The `abair.ie` Project
  - Unit Selection Synthesis of Irish Dialects

# How HMM Synthesis Works



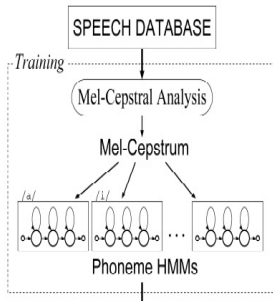


# Training HMMs



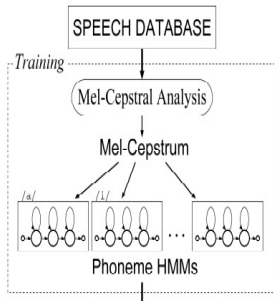
- Duration,  $f_0$  and spectral information is extracted from the recorded speech
- This information is stored as excitation and spectral parameters
- The parameters are modelled using Hidden Markov Models (HMM)

# Training HMMs



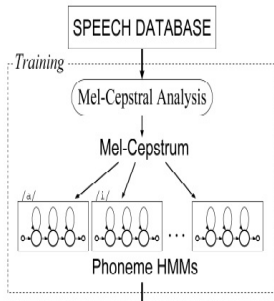
- Duration,  $f_0$  and spectral information is extracted from the recorded speech
- This information is stored as excitation and spectral parameters
- The parameters are modelled using Hidden Markov Models (HMM)

# Training HMMs



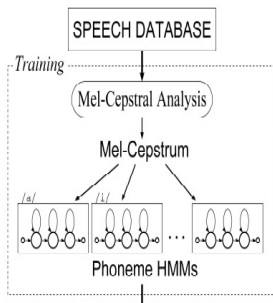
- Duration,  $f_0$  and spectral information is extracted from the recorded speech
- This information is stored as excitation and spectral parameters
- The parameters are modelled using Hidden Markov Models (HMM)

# Training HMMs



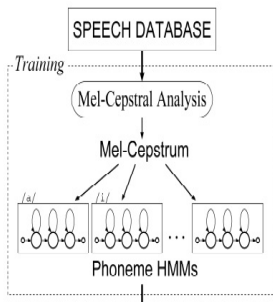
- Duration,  $f_0$  and spectral information is extracted from the recorded speech
- This information is stored as excitation and spectral parameters
- The parameters are modelled using Hidden Markov Models (HMM)

# Training HMMs



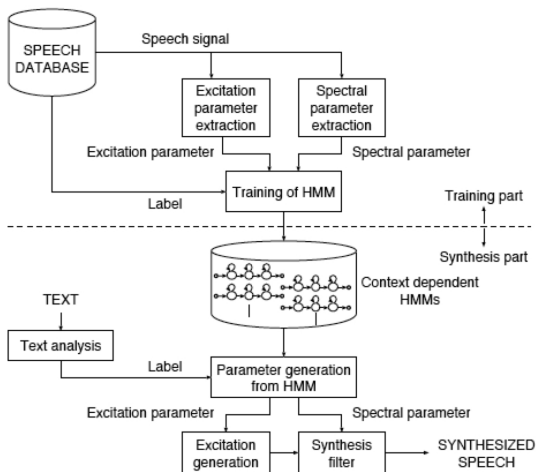
- Duration,  $f_0$  and spectral information is extracted from the recorded speech
- This information is stored as excitation and spectral parameters
- The parameters are modelled using Hidden Markov Models (HMM)

# Training HMMs

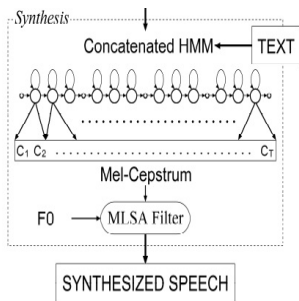


- Duration,  $f_0$  and spectral information is extracted from the recorded speech
- This information is stored as excitation and spectral parameters
- The parameters are modelled using Hidden Markov Models (HMM)

# How HMM Synthesis Works



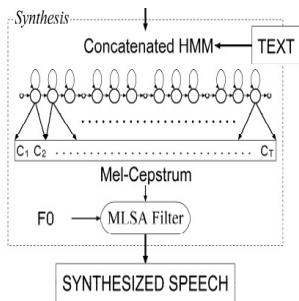
# Synthesising HMMs



- The target utterance HMM is constructed as a sequence of unit HMMs
- The excitation and spectral parameters are generated from the HMM
- The speech waveform is generated from the speech parameters using a filter

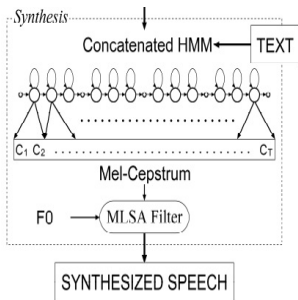


# Synthesising HMMs



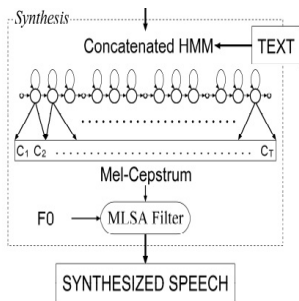
- The target utterance HMM is constructed as a sequence of unit HMMs
- The excitation and spectral parameters are generated from the HMM
- The speech waveform is generated from the speech parameters using a filter

# Synthesising HMMs



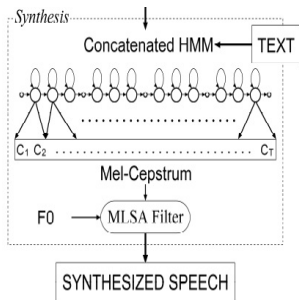
- The target utterance HMM is constructed as a sequence of unit HMMs
- The excitation and spectral parameters are generated from the HMM
- The speech waveform is generated from the speech parameters using a filter

# Synthesising HMMs



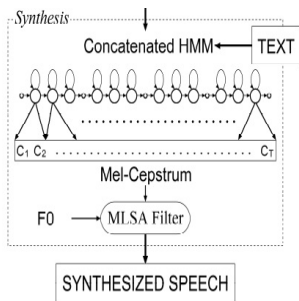
- The target utterance HMM is constructed as a sequence of unit HMMs
- The excitation and spectral parameters are generated from the HMM
- The speech waveform is generated from the speech parameters using a filter

# Synthesising HMMs



- The target utterance HMM is constructed as a sequence of unit HMMs
- The excitation and spectral parameters are generated from the HMM
- The speech waveform is generated from the speech parameters using a filter

# Synthesising HMMs



- The target utterance HMM is constructed as a sequence of unit HMMs
- The excitation and spectral parameters are generated from the HMM
- The speech waveform is generated from the speech parameters using a filter

# Pros and Cons of HMM Synthesis

## Pros

- Fast, accurate speech synthesis
- Requires only a small amount of memory
- Voices can be manipulated for voice disguise and emotional speech synthesis

## Cons

- Heavy signal processing gives the voice a buzzy quality
- This method of synthesis is still being developed

# Pros and Cons of HMM Synthesis

## Pros

- **Fast, accurate speech synthesis**
- Requires only a small amount of memory
- Voices can be manipulated for voice disguise and emotional speech synthesis

## Cons

- Heavy signal processing gives the voice a buzzy quality
- This method of synthesis is still being developed

# Pros and Cons of HMM Synthesis

## Pros

- Fast, accurate speech synthesis
- Requires only a small amount of memory
- Voices can be manipulated for voice disguise and emotional speech synthesis

## Cons

- Heavy signal processing gives the voice a buzzy quality
- This method of synthesis is still being developed



# Pros and Cons of HMM Synthesis

## Pros

- Fast, accurate speech synthesis
- Requires only a small amount of memory
- Voices can be manipulated for voice disguise and emotional speech synthesis

## Cons

- Heavy signal processing gives the voice a buzzy quality
- This method of synthesis is still being developed

# Pros and Cons of HMM Synthesis

## Pros

- Fast, accurate speech synthesis
- Requires only a small amount of memory
- Voices can be manipulated for voice disguise and emotional speech synthesis

## Cons

- Heavy signal processing gives the voice a buzzy quality
- This method of synthesis is still being developed

# Pros and Cons of HMM Synthesis

## Pros

- Fast, accurate speech synthesis
- Requires only a small amount of memory
- Voices can be manipulated for voice disguise and emotional speech synthesis

## Cons

- Heavy signal processing gives the voice a buzzy quality
- This method of synthesis is still being developed

# Pros and Cons of HMM Synthesis

## Pros

- Fast, accurate speech synthesis
- Requires only a small amount of memory
- Voices can be manipulated for voice disguise and emotional speech synthesis

## Cons

- Heavy signal processing gives the voice a buzzy quality
- This method of synthesis is still being developed

# Outline

- 1 The Irish Language
  - Overview of the Irish Language
  - Applications of Minority Language Synthesis
- 2 Methods of Concatenative Speech Synthesis
  - Pre-Requisites for Corpus-Based Synthesis
  - Unit Selection Synthesis
  - HMM Synthesis
- 3 The `abair.ie` Project
  - Unit Selection Synthesis of Irish Dialects

# Building the synthesiser

## Building the Corpus

- Gathering out of copyright, dialect specific material for recording
- Recording, aligning and annotating the corpus

## The Lexicon

- Creating a dictionary of Irish words and pronunciations

## Letter-to-Sound Rules

- Creating a list of grapheme-to-phoneme rules to cater for out-of-lexicon words

# Building the synthesiser

## Building the Corpus

- Gathering out of copyright, dialect specific material for recording
- Recording, aligning and annotating the corpus

## The Lexicon

- Creating a dictionary of Irish words and pronunciations

## Letter-to-Sound Rules

- Creating a list of grapheme-to-phoneme rules to cater for out-of-lexicon words

# Building the synthesiser

## Building the Corpus

- Gathering out of copyright, dialect specific material for recording
- Recording, aligning and annotating the corpus

## The Lexicon

- Creating a dictionary of Irish words and pronunciations

## Letter-to-Sound Rules

- Creating a list of grapheme-to-phoneme rules to cater for out-of-lexicon words



# Building the synthesiser

## Building the Corpus

- Gathering out of copyright, dialect specific material for recording
- Recording, aligning and annotating the corpus

## The Lexicon

- Creating a dictionary of Irish words and pronunciations

## Letter-to-Sound Rules

- Creating a list of grapheme-to-phoneme rules to cater for out-of-lexicon words

# Building the synthesiser

## Building the Corpus

- Gathering out of copyright, dialect specific material for recording
- Recording, aligning and annotating the corpus

## The Lexicon

- Creating a dictionary of Irish words and pronunciations

## Letter-to-Sound Rules

- Creating a list of grapheme-to-phoneme rules to cater for out-of-lexicon words

# Building the synthesiser

## Building the Corpus

- Gathering out of copyright, dialect specific material for recording
- Recording, aligning and annotating the corpus

## The Lexicon

- Creating a dictionary of Irish words and pronunciations

## Letter-to-Sound Rules

- Creating a list of grapheme-to-phoneme rules to cater for out-of-lexicon words

# Building the synthesiser

## Building the Corpus

- Gathering out of copyright, dialect specific material for recording
- Recording, aligning and annotating the corpus

## The Lexicon

- Creating a dictionary of Irish words and pronunciations

## Letter-to-Sound Rules

- Creating a list of grapheme-to-phoneme rules to cater for out-of-lexicon words

# Building the tools

## Normalisation Tool

- Changes dates, times and other notation into a string of letters
- Very complex in Irish

## Phonetisation Tool

- Changes normal text into a phonetic computer friendly sequence of units
- The corpus is labelled in this way so the units can then be found during synthesis

# Building the tools

## Normalisation Tool

- Changes dates, times and other notation into a string of letters
- Very complex in Irish

## Phonetisation Tool

- Changes normal text into a phonetic computer friendly sequence of units
- The corpus is labelled in this way so the units can then be found during synthesis

# Building the tools

## Normalisation Tool

- Changes dates, times and other notation into a string of letters
- Very complex in Irish

## Phonetisation Tool

- Changes normal text into a phonetic computer friendly sequence of units
- The corpus is labelled in this way so the units can then be found during synthesis

## Building the tools

### Normalisation Tool

- Changes dates, times and other notation into a string of letters
- Very complex in Irish

### Phonetisation Tool

- Changes normal text into a phonetic computer friendly sequence of units
- The corpus is labelled in this way so the units can then be found during synthesis



# Building the tools

## Normalisation Tool

- Changes dates, times and other notation into a string of letters
- Very complex in Irish

## Phonetisation Tool

- Changes normal text into a phonetic computer friendly sequence of units
- The corpus is labelled in this way so the units can then be found during synthesis

## Building the tools

### Normalisation Tool

- Changes dates, times and other notation into a string of letters
- Very complex in Irish

### Phonetisation Tool

- Changes normal text into a phonetic computer friendly sequence of units
- The corpus is labelled in this way so the units can then be found during synthesis

# `abair.ie` Interface

The screenshot shows the `abair.ie` web interface in a browser window. The page header includes the site name and navigation links. The main content area features a search box, a language selection dropdown (set to 'English'), and a text input field. Below the input field is a 'Synthesise' button and a 'Listen' button. A 'Feedback form' is also visible. The left sidebar contains a 'Suite of tools' menu with options like 'Help us', 'Background', 'What is synthesis?', 'People', 'Contact', and 'FAQ'. The main content area includes a 'Say it!' section with a text input field containing 'Beath na bairne í a labairt'. The interface is annotated with several callouts: 'Search box' points to the search input; 'Choice of interface language' points to the language dropdown; '1. Enter text' points to the text input field; '2. Click' points to the 'Synthesise' button; '3. Listen' points to the 'Listen' button; '4. Save (optional)' points to the 'Feedback form'; and 'Suite of tools' points to the left sidebar menu.

Search box

Choice of interface language

Suite of tools

1. Enter text

2. Click

3. Listen

4. Save (optional)

Feedback form

# Summary

- Overview of **the Irish language**, and the importance of synthesising minority languages.
- Explanation of **unit selection** and **Hidden-Markov Model based synthesis**.
- The steps taken to create the first **Irish language speech synthesiser**.
- Outlook
  - Creating HMM-based voices for screen readers
  - Create some unit-selection/HMM hybrid for controllable speech synthesis

# Summary

- Overview of **the Irish language**, and the importance of synthesising minority languages.
- Explanation of **unit selection** and **Hidden-Markov Model based synthesis**.
- The steps taken to create the first **Irish language speech synthesiser**.
- Outlook
  - Creating HMM-based voices for screen readers
  - Create some unit-selection/HMM hybrid for controllable speech synthesis

# Summary

- Overview of **the Irish language**, and the importance of synthesising minority languages.
- Explanation of **unit selection** and **Hidden-Markov Model based synthesis**.
- The steps taken to create the first **Irish language speech synthesiser**.
- Outlook
  - Creating HMM-based voices for screen readers
  - Create some unit-selection/HMM hybrid for controllable speech synthesis

# Summary

- Overview of **the Irish language**, and the importance of synthesising minority languages.
- Explanation of **unit selection** and **Hidden-Markov Model based synthesis**.
- The steps taken to create the first **Irish language speech synthesiser**.
- Outlook
  - Creating HMM-based voices for screen readers
  - Create some unit-selection/HMM hybrid for controllable speech synthesis

# Summary

- Overview of **the Irish language**, and the importance of synthesising minority languages.
- Explanation of **unit selection** and **Hidden-Markov Model based synthesis**.
- The steps taken to create the first **Irish language speech synthesiser**.
- Outlook
  - Creating HMM-based voices for screen readers
  - Create some unit-selection/HMM hybrid for controllable speech synthesis



# Summary

- Overview of **the Irish language**, and the importance of synthesising minority languages.
- Explanation of **unit selection** and **Hidden-Markov Model based synthesis**.
- The steps taken to create the first **Irish language speech synthesiser**.
- Outlook
  - Creating HMM-based voices for screen readers
  - Create some unit-selection/HMM hybrid for controllable speech synthesis