Oxford University Phonetics Laboratory

UNIVERSITY OF OXFORD

Patterns of durational variation in British dialects

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Anastassia Loukina and Greg Kochanski

13 September 2010

This talk was presented at PAC workshop in Montpellier, France on 13 September 2010.

The reference is:

Loukina, A., Kochanski, G. (2010) “Patterns of durational variation in British dialect”. Paper presented at PAC workshop in Montpellier, France on 13 September 2010.

Introduction: rhythmic typology

- Stress-timed vs. Syllable-timed languages (Abercrombie, Pike)
- No evidence for isochrony (Dauer 1980, Roach 1982 etc.).
- Rhythm classes based on measures of duration (Ramus 1999, Grabe 2002)
- Large overlap in values of rhythm measure. (Keane 2006, Loukina et al. 2009)
- The experimental evidence for the existence of rhythm classes is inconclusive.

When we listen to somebody speaking a different language or even a different dialect, we often notice that their speech has different 'rhythm'. In mid 1960-s Abercrombie and Pike suggested that all languages could be divided into several rhythm classes. They argued that in languages like English there is a tendency for equal spacing between stressed syllables. Such languages with a “morse-code” rhythm have been called “stress-timed”. On the other hand, in languages like French, there is a tendency for equal distance between the syllables. This, supposedly, would create the impression of a “machine-gun” rhythm. These languages in turn have been called “syllable-timed”.

Unfortunately, subsequent experimental studies failed to find any evidence for isochrony in the actual phonetic data. However, the hypothesis that there are distinct rhythm classes was intuitively appealing and appeared to be somewhat supported by the perceptual data, so it was recast in terms of variation in duration of vocalic and consonantal intervals.

It was argued that the so-called 'syllable-timed' languages would show less vowels reduction and simpler syllable structure leading to less variation in duration of vowels and consonants. On the other hand, 'stress-timed' languages would show greater variation due to vowel reduction and more complex syllable structure (cf. Dauer 1980). In 1999 Ramus and colleagues suggested a set of measures aimed to capture variation between different languages. Almost at the same time Francis Nolan suggested a similar measure to account for variation in several varieties of English. The methodology has become hugely popular and a lot of new measures have been added to the original set.

Despite the popularity of these measures, very soon it has become clear that these measures are not robust to individual variation and differences between speakers could sometimes exceed differences between languages.

Rhythm in British English

- English: stress-timed
- Perceived difference in rhythm between dialects.
- Ferragne(2004, 2008): overlap in rhythm measures between Britsh dialects.
- Some varieties of English described as syllable-timed: Punjabi, West Indian, Welsh speakers, L2 English.

English has been traditionally considered as a prototypical “stress-timed” language, yet there is a widespread perception that different varieties of English have different “rhythm”. Some of the aforementioned rhythm measures have been applied to the dialects of English and, as in case of languages, it was found that there is a large overlap between different dialects.

Of course, this does not contradict the rhythm class hypothesis: if dialects can be split into several classes, one would expect that only dialects which belong to different classes would be differentiated by durational metrics. Dialects within the same class should indeed show similar properties.

In the context of the British Isles, it has been often claimed that such varieties as Punjabi English, Welsh English or West Indian English may show tendency towards syllable-timing and thus differ from other more “stress-timed” dialects.

Do traditional rhythm classes capture variation in rhythmic patterns across British English or do we need more dimensions?

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 4

In this talk we address the question whether the traditional division into rhythm classes allows capturing variation in rhythm among British dialects or a more complex system is needed.

Methodology

Step 1: Select the data.

Step 2: Segment the data into vowel-like and consonant-like intervals.

Step 3: Compute the rhythm measures.

Step 4: Compare the dialects.

We will start by describing the methodology we used: the data, the segmentation, the rhythm measures and the methods we used to compare the dialects.

Data: the IViE corpus

- **I**ntonational **V**ariation **i**n **E**nglish
- Recordings in several urban locations of British isles
- Collected in 1997-2000 by E. Grabe, F. Nolan, B. Post and K.Farrar.
- Speakers:
  - Age: 16 years old
  - Recorded in their schools

www.phon.ox.ac.uk/IViE

UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 6

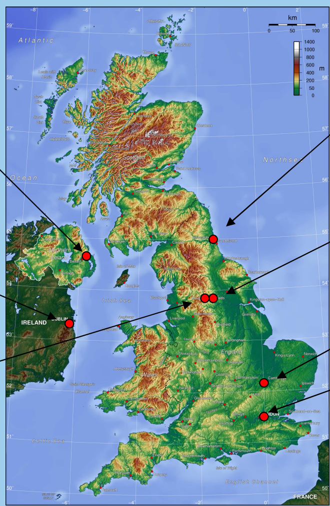
This study is based on the IViE corpus collected in 1997-2000 by F. Nolan, E. Grabe, B. Post and K. Farrarn and available to download at our website. The corpus was recorded in schools in several urban location in Britain and Ireland and contains read and spontaneous speech of adolescents from these areas.

IViE: The dialects used in this study

Belfast

Dublin

Bradford
(bilingual residents of
Punjabi decent)



Newcastle

Leeds

Cambridge

London
(monolingual residents of
Jamaican decent)

UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

We used the recordings from 7 dialects: Cambridge, Leeds, Newcastle, Belfast, Dublin, the speech of bilingual Punjabi residents of Dublin and London residents of Jamaican decent.

Data: the corpus used for this study

- 12 speakers for each dialect
  6M+6F (7F+5M for London)
- 6 texts per speaker
  Reading (5 parts) and re-telling of Cinderella story
- **Total: 465 short texts (ca. 180 syllables each).**

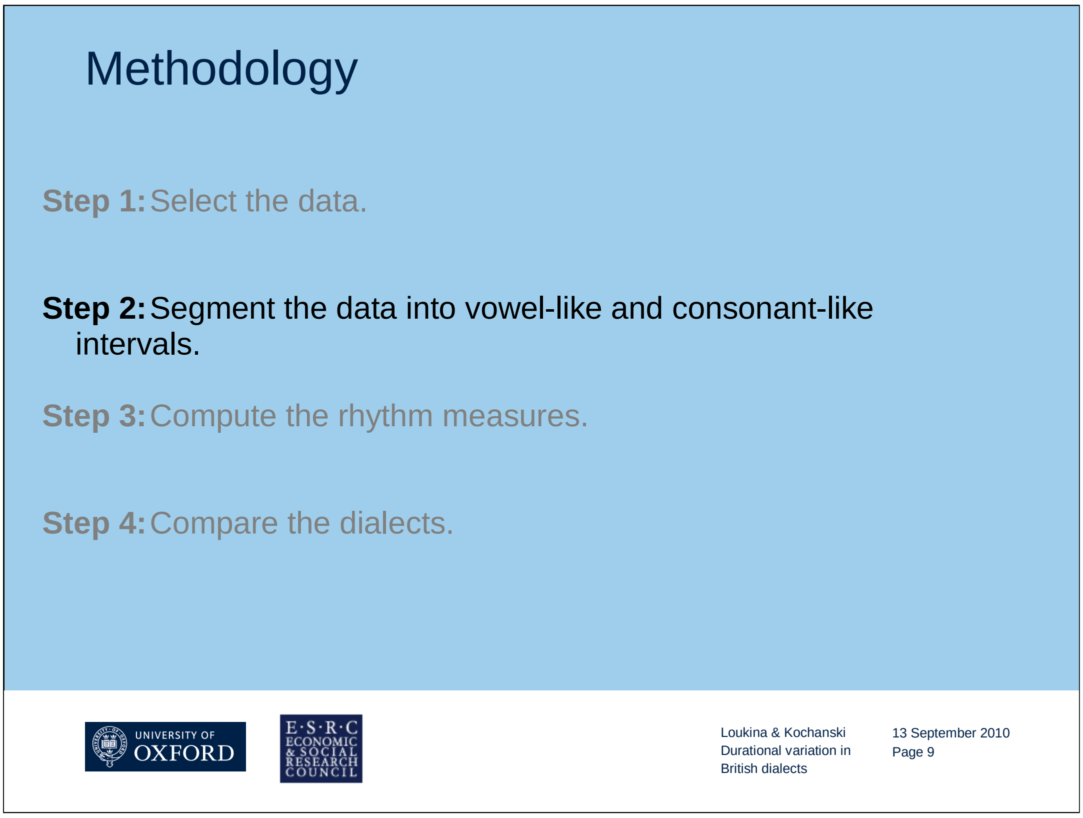
UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 8

Although the IViE corpus contains samples of different styles of speech, we only included reading and re-telling of Cinderella story recorded from 12 speakers from each dialect. After we discarded the texts with low quality of the recordings, our subcorpus contained 465 texts.



The computation of rhythm metrics is based on consonantal and vocalic intervals. We will now explain how we segmented the data into such intervals.

Segmentation

- Automatic segmentation into Consonant-like and Vowel-like segments using HTK-Toolkit.
- Why automatic segmentation?
  - ✓Consistency
  - ✓Strictly acoustic criteria
  - ✓Independent of phonological interpretation
    - Syllabic consonants, glides, devoiced vowels
  - ✓Large corpora of data

UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 10

We used HTK-toolkit (http://htk.eng.cam.ac.uk) to segment the data into vowel-like and consonant-like intervals. First, we created statistical models of 'vowels' and 'consonants' based on manual segmentation of a subset of data and then used these models to segment the remaining data.

We have chosen to use automatic segmentation for several reasons.

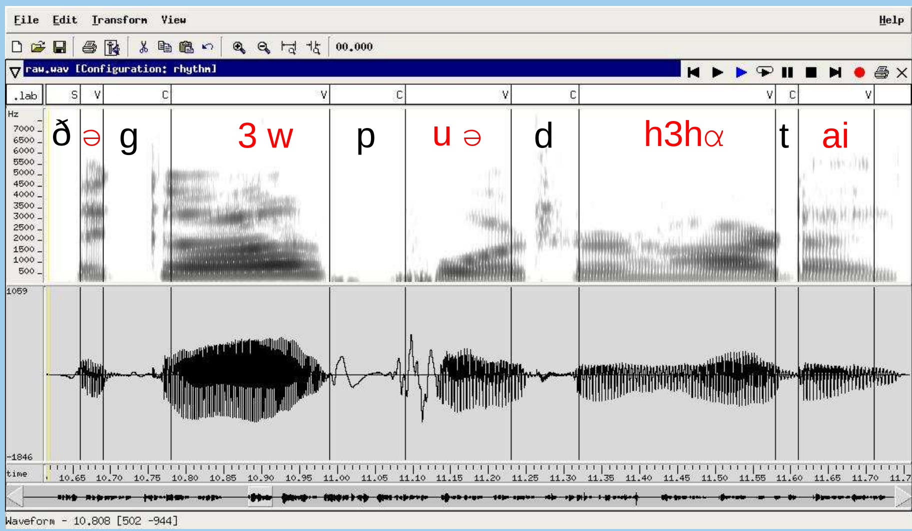
First, manual segmentation is a subjective and often error-prone process. Automatic segmentation offers more consistent approach.

Second, manual labellers may be influenced by their pre-existing knowledge about the language or dialect, while an automatic segmentation is based on strictly acoustic criteria.

As a result automatic segmentation is independent of possible differences in phonological interpretation, especially in such ambiguous cases as devoiced vowels, syllabic consonants or glides.

Last but not least automatic segmentation allows using large corpora of data and as we will show in this talk this is an absolute necessity in such studies.

Segmentation



The girl poured her heart out [Belfast]

Loukina & Kochanski
Durational variation in
British dialects

On this slide you can see an example of segmented data.

Segmentation: evaluation

- 10 minutes of manually segmented data.
- Cohen's Kappa:

  A measure of inter-rater agreement which takes into account agreement by chance.

  κ<0 – agreement no better than chance

  κ>0.75 - excellent agreement
- κ=0.64 – fair to good agreement.

UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

To evaluate the automatic segmentation, we have asked 2 trained phoneticians to segment about 10 minutes of data and used Cohen's kappa to rate the agreement between human and automatic segmentation (see Loukina et al., submitted, for further details).

We found good agreement.

Methodology

Step 1: Select the data.

Step 2: Segment the data into vowel-like and consonant-like intervals.

Step 3: Compute the rhythm measures.

Step 4: Compare the dialects.

PVI-like measures for vocalic and consonantal interval

- **CrPVI** – Raw consonantal pairwise variability index
- **VnPVI** – Normalized vocalic pairwise variability index
- **CnPVI** – Normalized consonantal PVI
- **med_CrPVI** – median CrPVI
- **med_VnPVI** – median VnPVI

Here you can see a list of measures that we have computed. For references see our Interspeech paper (Loukina et al. 2009)) available at our website www.phon.ox.ac.uk/speech_rhythm.

Global measures for vocalic and consonantal intervals

- **%V** – Percentage of vocalic intervals
- **Vdur/Cdur** – Ratio of vowel duration to consonant duration

Ratio

- **ΔV** – Std. deviation of vocalic intervals
- **ΔC** – Std. deviation of consonantal intervals
- **VarcoΔV** – ΔV/mean vocalic duration
- **VarcoΔC** – ΔC/mean consonantal duration

PVI-like measures for CV intervals

- **PVI-CV** – raw PVI of consonant+vowel group
- **VI** – PVI normalized by average duration
- **YARD** – normalized by z-scores
- **nCVPVI** – Normalized PVI of C+V

How many rhythm measure?

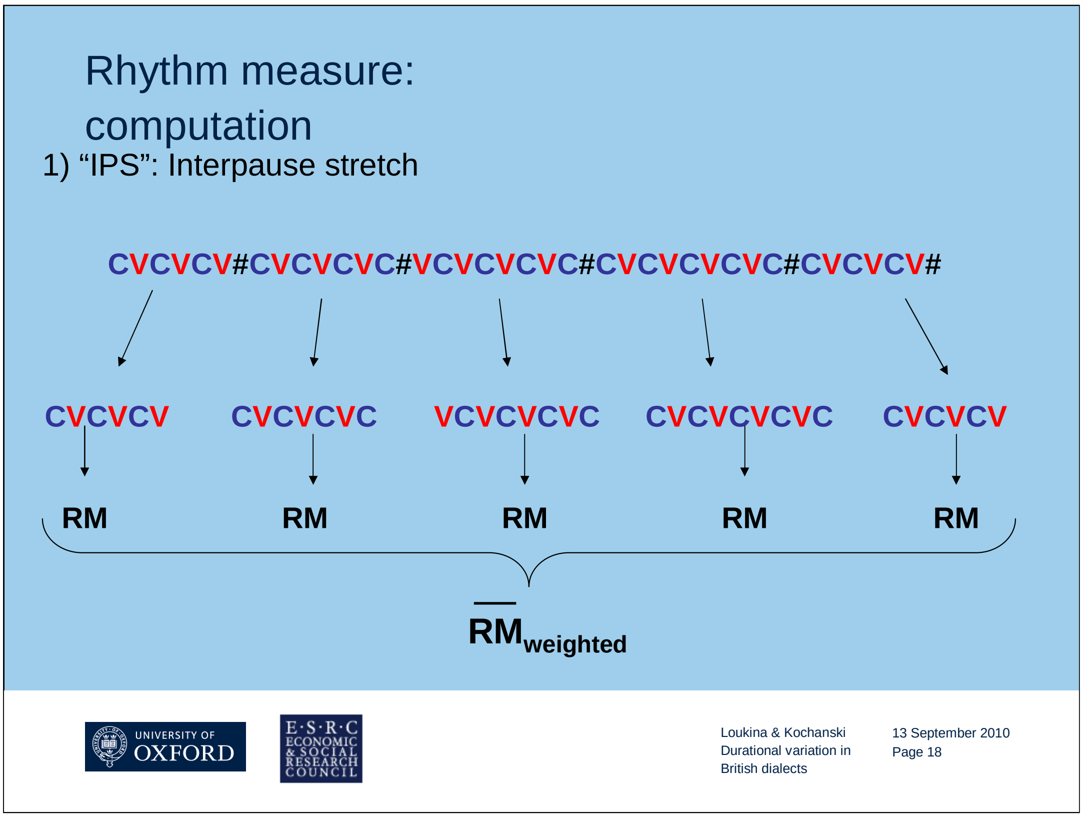
- 15 single measures
- 105 combinations of two measures
- 455 combinations of three measures

Which of them best capture differences between the dialects?

Loukina & Kochanski
Durational variation in British dialects

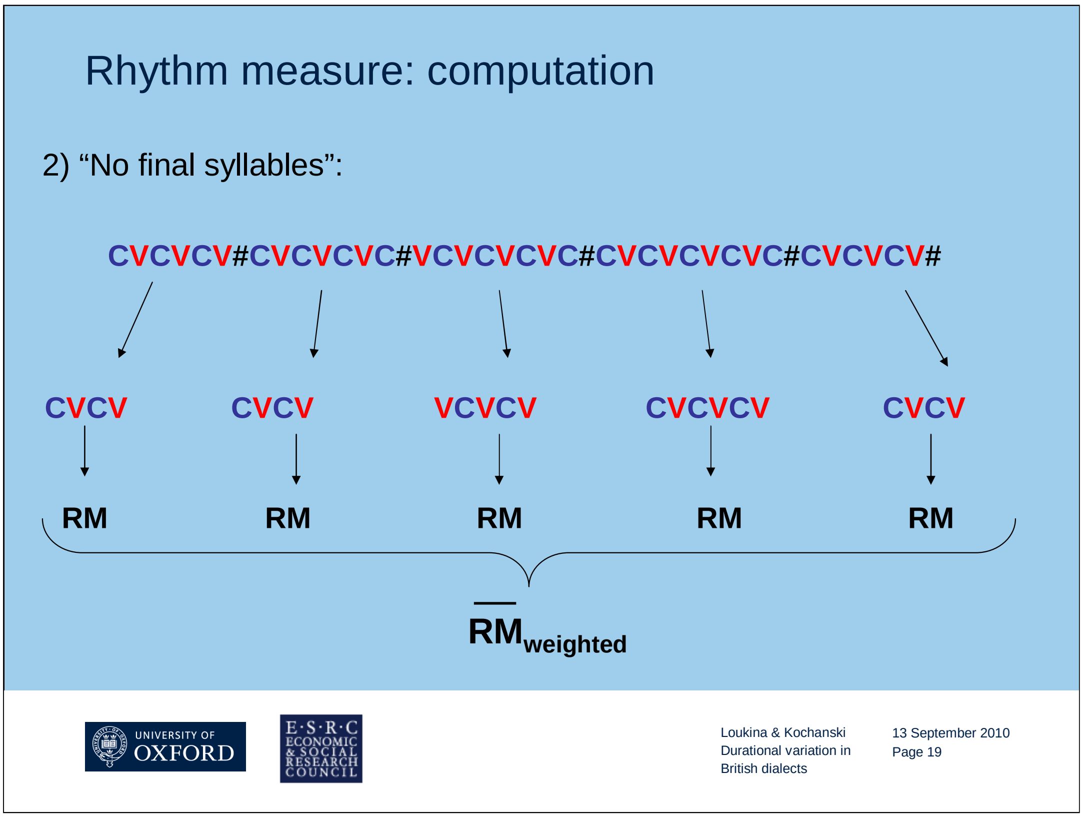
13 September 2010
Page 17

In total we computed 15 rhythm measures. Most previous studies were usually based on 2-3 measures and focused on most common combinations such as VnPVI+CrPVI or %V+△C. Yet, simple calculations would show that there are 105 possible combinations of two measures and 455 possible combinations of three measures. One of the questions we will address in this talk is which of these combinations give the best account of the variation in rhythm in 7 chosen dialects of English.



Previous research suggested three possible ways of computing the rhythm measures in texts which contain pauses.

First, rhythm measures can be computed for each inter-pause stretch and then averaged across the paragraph (we took an average value weighted by the duration of each inter-pause stretch).



In another approach, the final syllable of each inter-pause stretch may be omitted to avoid the effects of phrase-final lengthening,

Rhythm measure: computation

3) “No pauses”

CVCVCV#CVCVCVC#VCVCVCVC#CVCVCVCVC#CVCVCV#

CVCVCV CVCVCVC VCVCVCVC CVCVCVCVC CVCVCV

CVCVCVCVCVCVCVCVCVCVCCVCVCVCVCCVCVCV

RM

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 20

Last, rhythm measures can also be computed across the whole text without taking pauses into account.

Methodology

Step 1: Select the data.

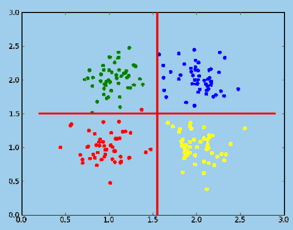
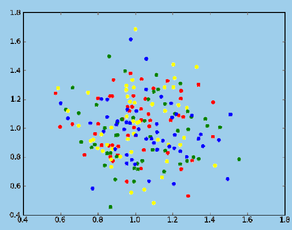
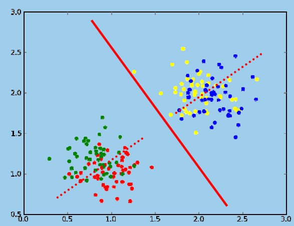
Step 2: Segment the data into vowel-like and consonant-like intervals.

Step 3: Compute the rhythm measures.

Step 4: Compare the dialects.

Classifier

How well can we identify a dialect based on a rhythm measure or a combination of several measures?

Distinct values	No difference	“Rhythm classes”
		
100%	Chance	≈75% consistent confusion patterns

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 22

To estimate the overlap in the values of RMs between different dialects, we used a machine-learning system which shows what is the probability of identifying the dialect of a paragraph based on the values of the RMs for the paragraph.

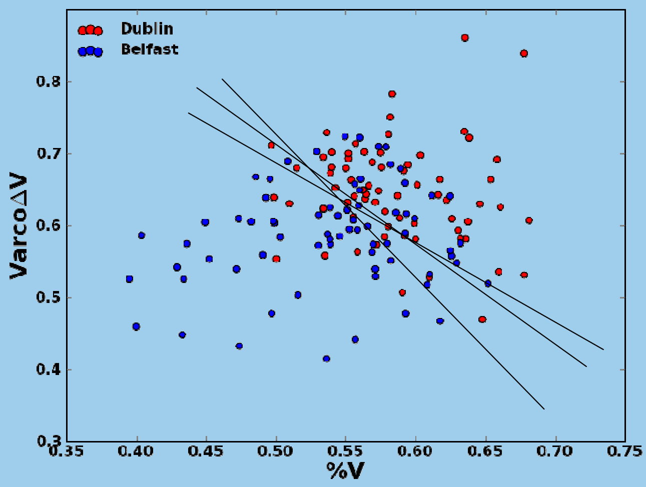
The classifier is an algorithm which estimates the boundaries between different dialects. If dialects were fully distinct as in the first figure, we would expect almost perfect identification. On the contrary if there were no difference in rhythm values between dialects, the identification would be at the chance level. Finally, if dialects formed several classes we would expect intermediate classification rates and consistent patterns of confusion between dialects which belong to the same class.

In the next few slides we will introduce the main steps in the application of classifier algorithm. For more technical description see Loukina et al. (submitted) available at our website.

The code has been released and is available at http://sourceforge.net/projects/speechresearch. Download packages g_classifiers-0.30.1.tar.gz and the libraries in gmisclib-0.67.9.tar.gz.

(These can also be downloaded from
http://www.phon.ox.ac.uk/files/releases/g_classifiers-0.30.1.tar.gz and
http://www.phon.ox.ac.uk/files/releases/gmisclib-0.67.9.tar.gz)

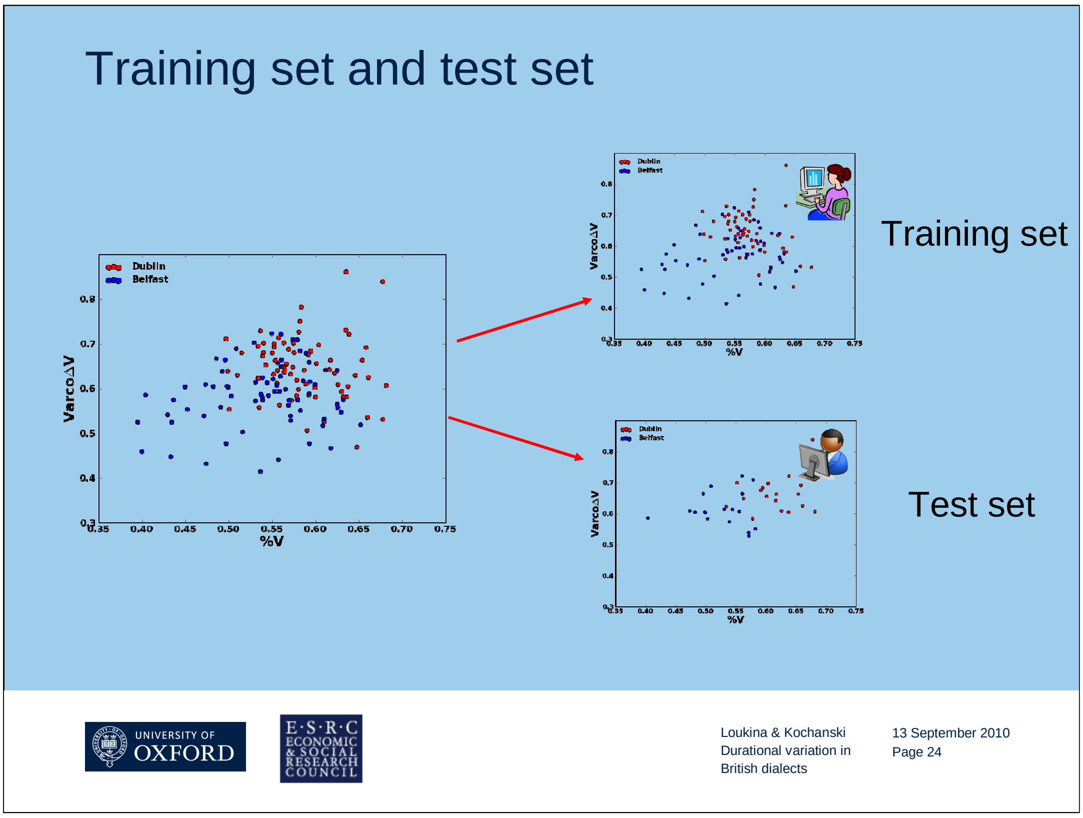
Classifier



For the sake of simplicity in these examples we will show the performance of the classifier which separates two dialects (Belfast and Dublin) based on a pair of measures (%V – Varco△V).

Each dot on the scatterplot shows the values of these two RMs for one paragraph in our corpus. As you can see, red dots tend to be in the upper right corner, blue dots tend to be in the lower left corner and there is also an area in the middle of the plot, where blue and red dots show an overlap. Based on visual analysis, one could draw several boundaries between the two classes.

So how do we decide which one is the best?

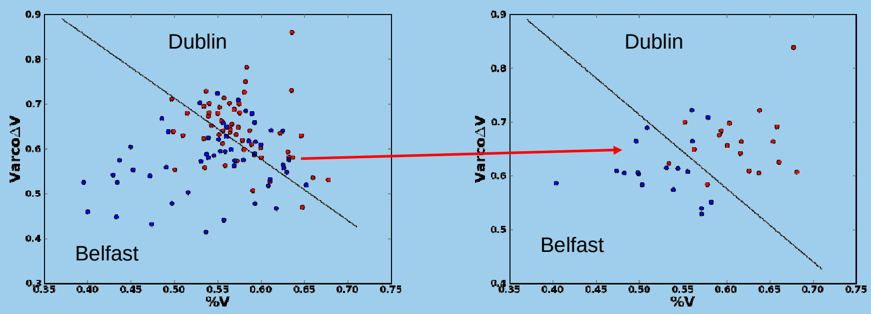


First we randomly split the data into what is usually called 'test set' and 'training set'. When doing this we made sure that the data from one speaker was always either in the training set or in the test set. This was done to separate individual and regional variation, that is we wanted the classifier to learn properties of a dialect and not of particular speakers. The question we asked was how well can you identify the dialect of speaker B once you learnt the properties of speaker A who is the speaker of the same dialect.

Classifier

Training set

Test set



86%

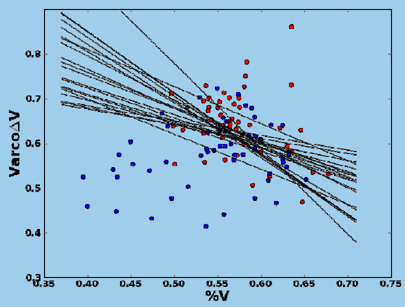
Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 25

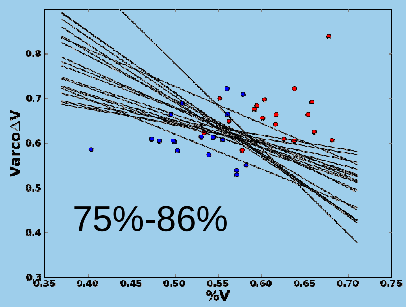
At the next step we found the linear equation that would best separate the dialects in the training set. This is basically a line that separates blue and red dots most successfully. We then applied the same equation to the test set and computed what share of the data was assigned to the correct class. As you can see, in this case 86% of the data were classified correctly.

Classifier

Training set



Test set



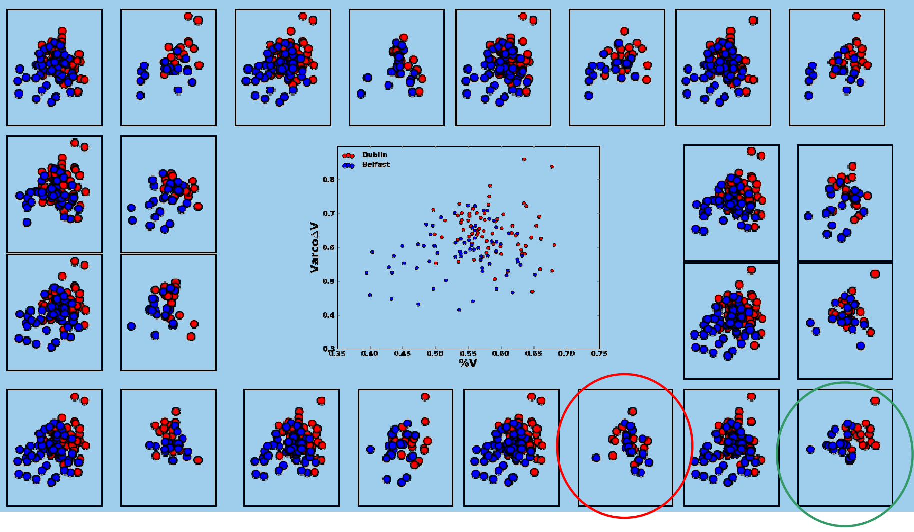
82%

20 classifiers with statistically equivalent performance

We then repeated the same process 20 times: we found 20 different equations (or lines) that showed best performance on the training set and applied them to the test set. We then computed the average of their performance on the test set. Note that although the classifiers showed equivalent performance on the training set, their performance on the test set varied between 75% and 86%. Therefore by taking an average we obtained a more accurate value which was less affected by a particularly lucky or particularly unlucky choice of equation.

Classifier

12 different combinations of test sets and training sets



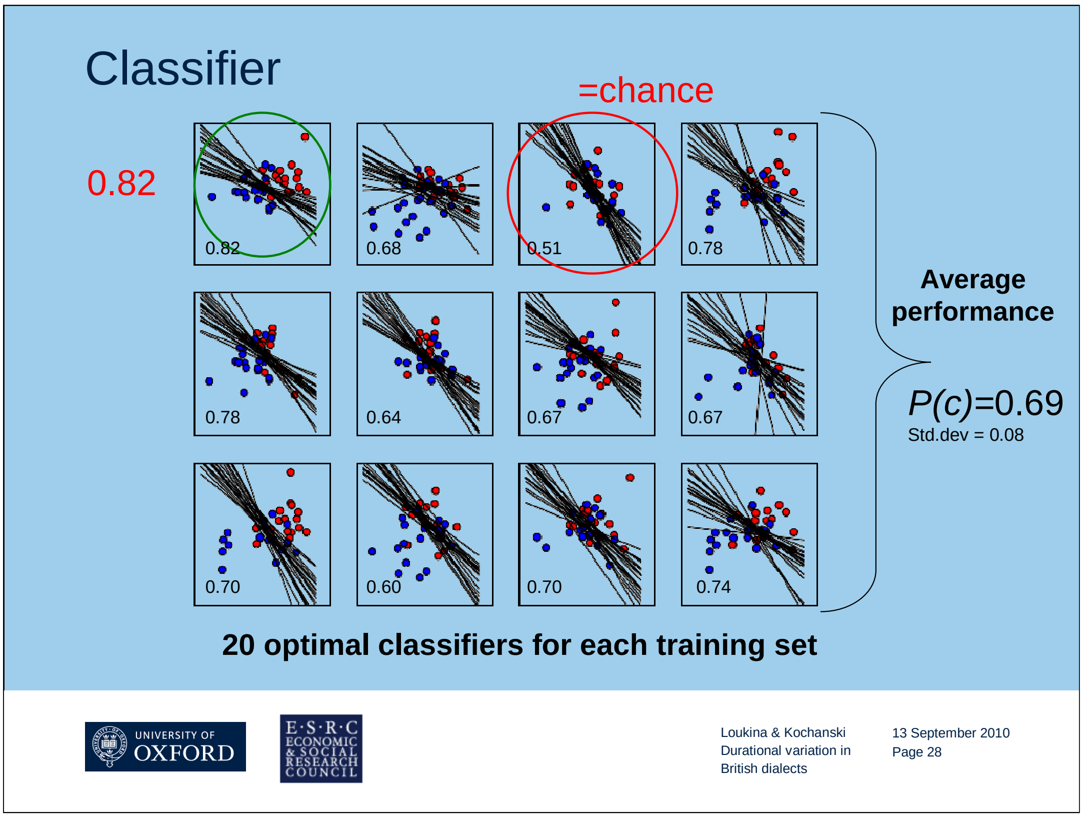
Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 27

Of course, the way we split the data between training set and test set could also affect our results. Therefore we repeated the same analysis with 12 different splits. One could think of these different combinations as different samples of data collected by different researches.

Note that the overlap between the values of rhythm measures varied between different splits: in the test set marked in green, blue and red dots show very clear pattern we have observed earlier: red dots in the upper right corner, left dots in the lower left corner. Yet, in the test set marked in red, both sets of dots completely overlap.

You may also note that in the test set in the lower left corner the position of red and blue dots is almost the reverse of what we see in the 'green' test set.



We then repeated our analysis on each of the combinations of training set and test set. As one would expect, in the 'green' set from the previous slide the probability of correct classification (P*(C)*) was 0.82, while in the 'red' set it was at the chance level. In other sets the P*(C)* varied with an average value of 0.69.

Remember, that we always place the data from the same speaker into either training set or test set. Therefore this slide shows that depending on the choice of speakers one may get very different results: the data from some speakers may be well separated, while others may show complete overlap. Furthermore, the differences in P*(C)* as high as 0.1-0.15 may be simply an artefact of data selection.

Lessons learnt:

- There is substantial individual variation: some speakers may consistently show very different patterns, while others can be almost indistinguishable.
- The choice of data and speakers may lead to substantial differences in classification rates.
- A difference as large as 10-15% can be a result of variation in the data.
- Resampling is crucial for reliable results.

UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 29

This slide summarizes the methodological conclusions.

These apply not only to production studies, but also to perception studies where the choice of data may lead to substantial variation in the results.

Results

- What is the best way to deal with pauses?
- How well can we separate the dialects and what measures are the best for this?
- Can we identify rhythmic classes within these dialects?

We will now present the results of our main study.

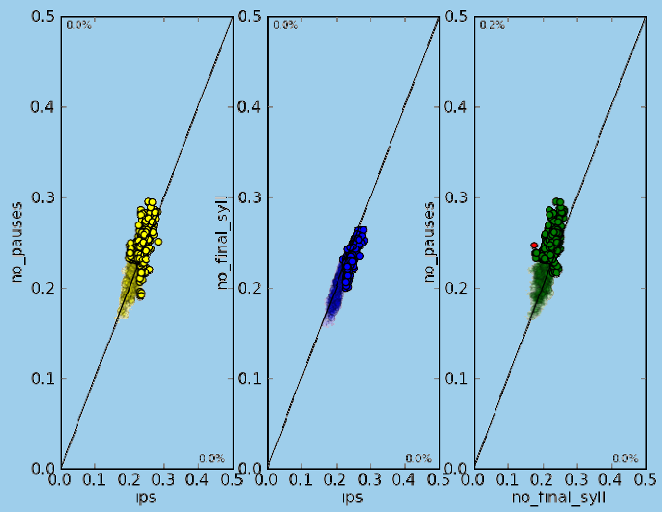
Analysis

- 15 single RMs x 3 ways of computation
- 105 combinations of two RMs x 3 ways of computation
- 455 combinations of three RMs x 3 ways of computation

1725 classifiers

I have already mentioned that there are many possible combinations of different rhythm measures and ways of computation. We have tried them all and will now present the results of these analyses.

Different ways of computation



No effect

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 32

First of all, different ways of computation had almost no effect on the result of classification. Each plot in this figure corresponds to a combination of two ways of computation. Each dot represents the results of classification based on the same combination of rhythm measures using two different ways of computation. The cases where the difference between two methods is significant are marked in red. As you can see, there was only one such case.

From the linguistic point of view, this means that the rhythmic differences between the dialects in this corpus are not restricted to different patterns of final lengthening on the last syllable. Otherwise, we would have seen different results for RMs computed with or without final syllable.

Results

- Is there the best method of computation?
- How well can we separate the dialects and what measures are the best for this?
- Can we identify rhythmic classes within these dialects?

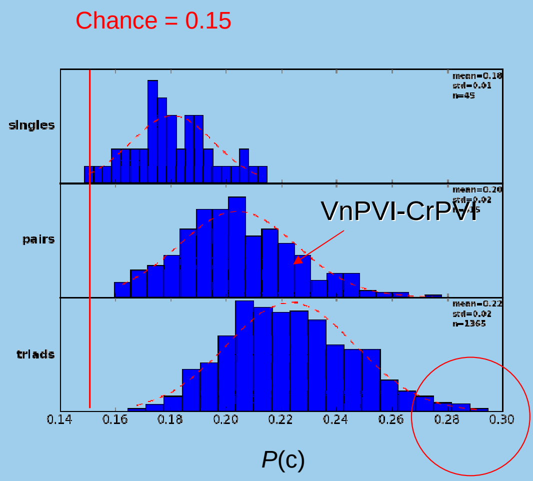
UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Our second question was how well different combinations of measures allow separating different dialects.

Results for duration

- RMs can separate dialects above chance
- There is large overlap in values
- Maximum probability of correct identification: 0.3
- Several measures allow achieving better separation between dialects.



Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 34

The figure shows how reliably our classifiers could identify the dialect based on different combinations of RMs. Each histogram shows the distribution of probabilities of correct classification (*P*(C)) based on one, two and three measures. The red vertical line shows the chance level, that is the expected P(C) if there were no difference between the dialects.

The histogram shows the classification rates achieved by all possible combinations of RMs.

To give an example, a popular combination of VnPVI and CrPVI allowed correctly classifying 22%.

The main conclusions are listed to the left of the figure.

What measures perform best?

- Non-normalized consonantal measures such as (CrPVI, ΔC) allow identification above chance.
- An addition of ratio measures (%V or Vdur/Cdur) leads to significant improvement.
- **Many combinations of different measures lead to comparable identification rates.**

UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Loukina & Kochanski
Durational variation in British dialects

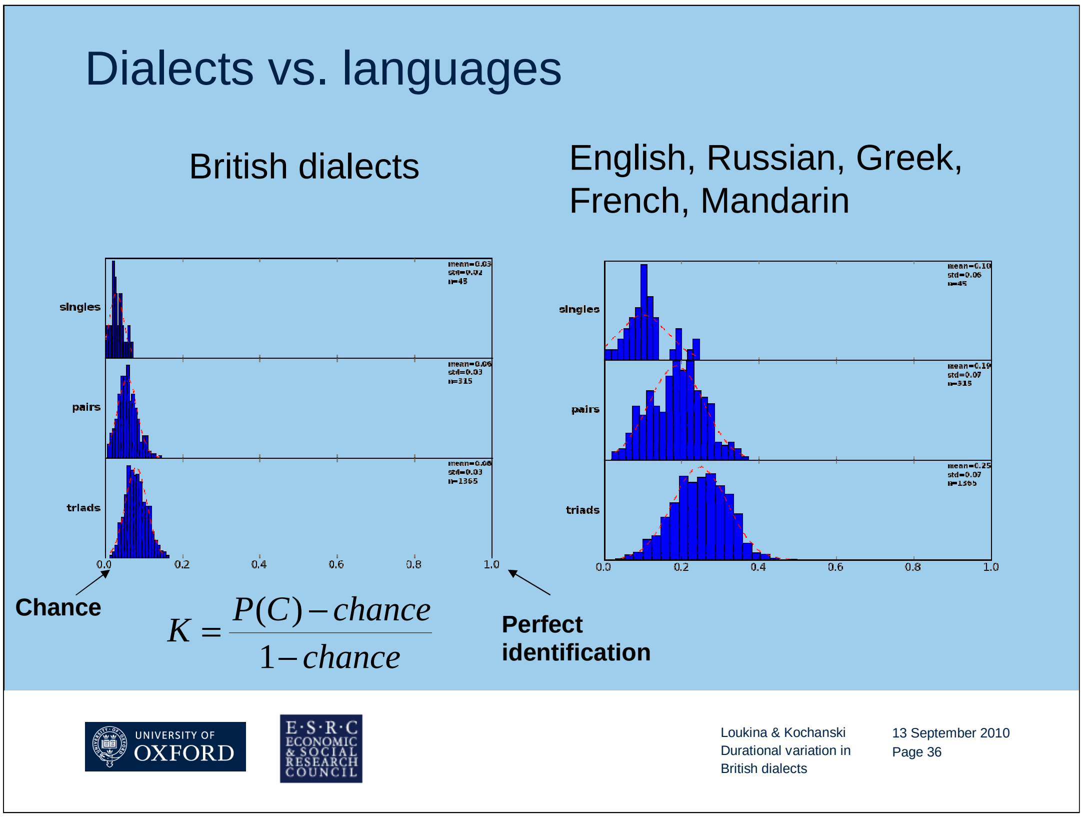
13 September 2010
Page 35

We found that for single measures only non-normalized consonantal measures allowed identification above chance. This points to potential differences in speech rate between speakers of different dialects which may or may not have linguistic significance.

The classification rates achieved based on consonantal measures could be significantly improved by adding ratio measures such as %V or Vdur/Cdur.

Furthermore, while no vowel-based measure separated the dialects on its own, a classifier based on a combination of several vowel-based measures achieved the identification rates similar to the consonantal and ratio based measures.

In general, there were many different combinations of measures that allowed comparable classification rates rather than one single best combination.



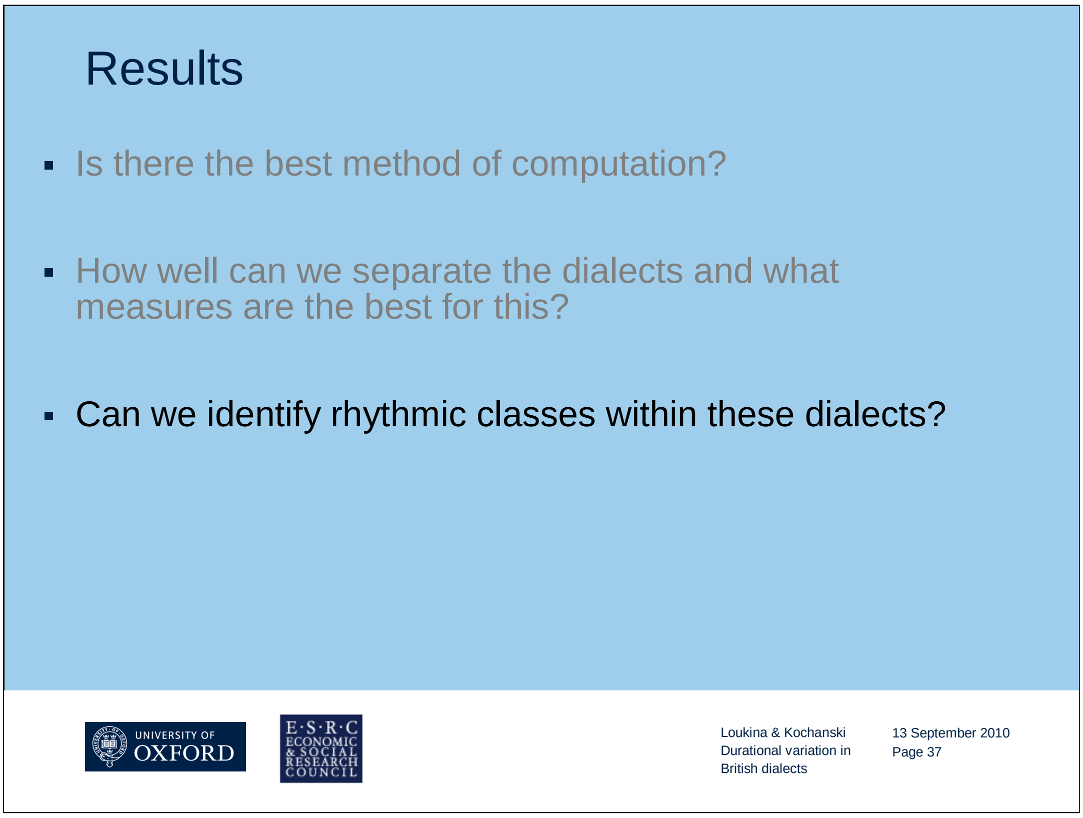
To put things into perspective, on this slide you can see how the classification rates for British dialects (left) compare to the classification rates for different languages (right). We performed a very similar analysis on a corpus of Cinderella stories read by native speakers of Russian, Southern British English, Greek, French and Taiwanese Mandarin (for full results of the analysis see other papers on our website).

Both histograms show classification rates achieved by single measures and combinations of several measures. Since chance values differed between the corpora, the X-axis shows *K*-value which allows for direct comparison. *K* varies between 0 for identification at chance level and 1 for perfect identification.

As you can see, even for different languages the classification rates are rather low, although higher than for British dialects. So it seems that rhythm is to a certain degree determined by the language and despite the differences between different dialects, we can speak about “English rhythm”.

Note that our observation that greater number of measures leads to better classification rates also holds for different languages.

Incidentally, the classification results for different languages agree well with the perceptual studies. When presented with low-pass filtered signal, people cannot separate languages at 100% and the reported identification rates are remarkably similar to the classification rates achieved by our classifiers.



Earlier I have mentioned that the overlap in values of the rhythm measures between different dialects would be expected if dialects were split between several rhythm classes. In this case we would expect consistent confusion patterns between dialects within the same class and good separation between dialects from different classes.

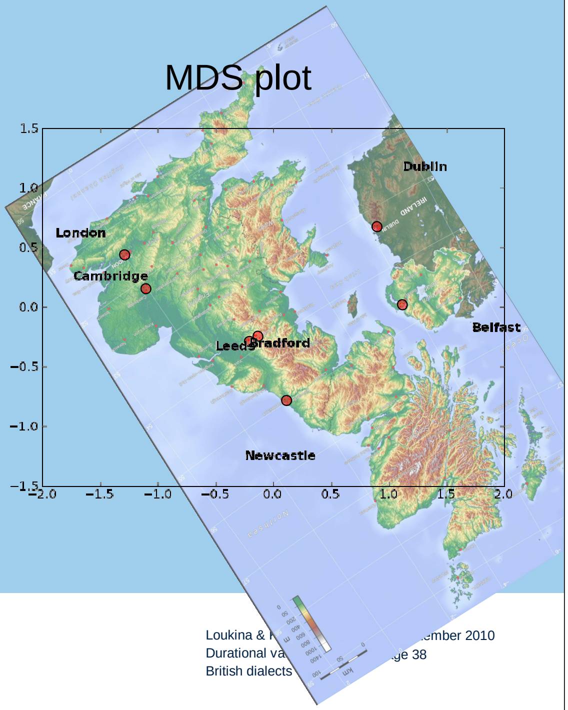
We used multidimensional scaling to study the grouping patterns between the dialects.

Multidimensional scaling: ALSCAL

Driving distances in km

	Belfast	Leeds	London	Cambridge	Dublin	Bradford	Newcastle
Belfast		456	749	670	166	286	344
Leeds	456		315	210	379	16	158
London	749	315		96	583	327	455
Cambridge	670	210	96		551	251	371
Dublin	166	379	583	551		371	543
Bradford	286	16	327	251	371		162
Newcastle	344	158	455	371	543	162	

MDS plot

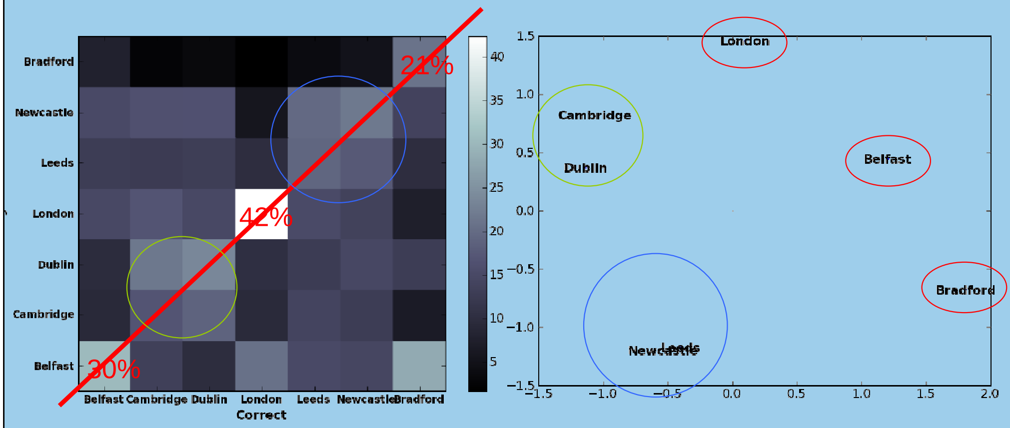


Multidimensional scaling allows to present dissimilarities and similarities in the data on a plot. Here is how it works.

If we take a table of driving distances between our cities (on the left) and perform ALSCAL MDS analysis we will get a picture presented on the right. As you can see the relative location of the cities (bold text) on what is called an MDS map is close to their geographical location (red dots), although the orientation of the map is random and bears no resemblance to traditional maps.

You may also note that our MDS Belfast and Dublin are particularly far from their geographical counterparts. This is of course because the table I used to construct the map contained driving rather than flying distances and there are only few places where one can cross over to Ireland which increases the driving distance between certain cities.

Multidimensional scaling: duration



Top 25% of all classifiers. Average *P*(C)=0.25

UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 39

And this is how we applied this technique to our data.

The figure on the left shows the averaged confusion matrix for best performing 25% of the classifiers.

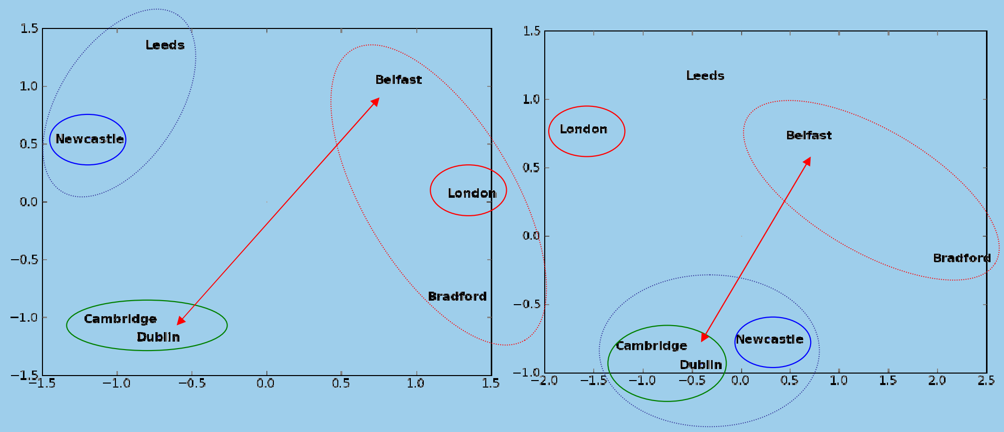
The columns show the correct dialect and the rows show the dialect assigned by the classifier.

That is top left cell shows what percentage of Belfast data was classified as coming from Bradford. Brighter squares correspond to higher percentage. The cells on the diagonal show the correctly classified data.

We have applied MDS analysis to the confusion matrix assuming that dialects that are confused more often should be closer on our imagined 'Rhythm map' that dialects that are always distinct. The result is shown on the right figure.

As you can see, London, Belfast and Bradford data stand apart and do not cluster with other dialects. Cambridge patterns together with Dublin and the two Northern dialects, Leeds and Newcastle also group together.

Duration: is there a single grouping?



Variation in Consonants   Vowels, CV and ratio

P(C)=0.25

UNIVERSITY OF OXFORD

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 40

Earlier I mentioned that different combinations of measures were equally successful in separating the dialects in our corpus. But were the confusion patterns consistent between different combinations?

The answer is **no**.

On this slide you see two MDS plots: the one on the left is based on the consonantal measures and the one on the right on non-consonantal measures. That is the right graph shows which dialects are more similar in patterns of variation of consonant duration and which dialects are more similar when it comes to the duration of vowels or CV sequences.

As you can see, Cambridge and Dublin are close on both graphs. Newcastle is more similar to Leeds in terms of variation in consonants, but closer to Dublin and Cambridge in terms of variation in vowels. Bradford, Belfast and London always stand apart but their location relative to each other and other dialects varies.

Noteworthy, Belfast and Dublin are far apart on both graphs, that is patterns of variation in duration are consistently different in these two dialects.

Duration: conclusions

- There is a large overlap in values of rhythm measures between different dialects.
- Cambridge and Dublin speakers tend to show similar patterns of variation.
- They are consistently different from Bradford (Punjabi), London (Jamaica) and Belfast speakers.
- Newcastle and Leeds show similar patterns of variation in consonants, but different patterns of variation in vowels.
- Variation in duration is multidimensional.

This slide summarizes the results for duration.

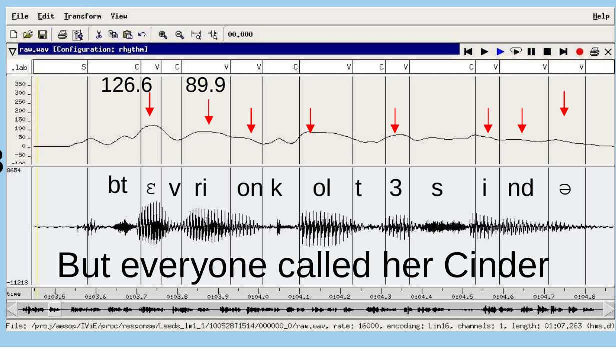
“LLog”: Differences in loudness

- Find maximum loudness of each ‘V’ label
- Compute the ratio of consecutive ‘V’ labels
- Take the absolute values of the *log* of that ratio.

  Values close to 0 indicate small difference.
- Compute mean value across the text

$$\frac{126.6}{89.9} = 1.4$$

$$|\ln(1.4)| = 2.3$$



Of course, the perceived differences in rhythm are unlikely to be attributed solely to duration. In fact, our results from another experiment show that rhythm may primarily rely on the rate of spectral change and loudness (see the Speech Prosody 2010 paper on our website). Therefore we have also looked into the variation in loudness.

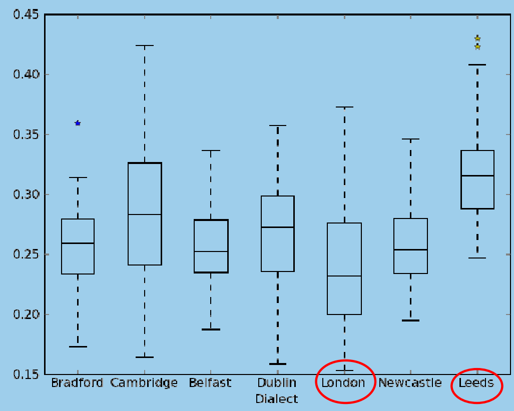
Rather than using raw intensity, we used a measure of estimated perceived loudness computed from the spectrum as described in Kochanski et al. 2005 (JASA, 118(2), 1038-1054).

We used these values to compute a new rhythm measure called Llog.

Llog is the average of absolute logarithms of ratios of maximum loudness values of adjacent vocalic segments. It is similar to PVI in that it measures how big is the contrast between two adjacent segments. Smaller values indicate little contrast, greater values correspond to substantial difference in loudness.

Loudness: classifier

- *P*(C) = 0.24 – similar to durational measures
- Leeds: greater differences between adjacent vowels. 63% of data correctly identified.
- London: smaller differences. 47% of data correctly identified.



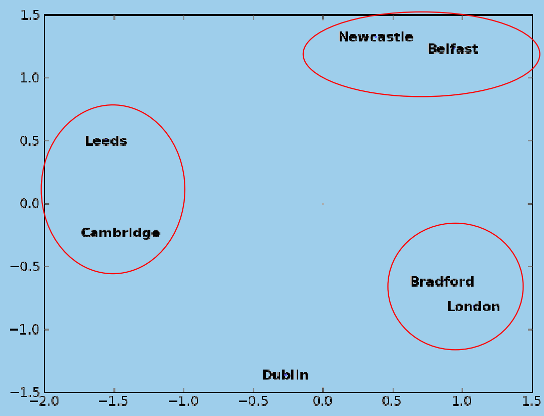
Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 43

We found that the dialects in our corpus differed in Llog, although as one would expect there was an overlap in values. A classifier based on Llog achieve *P*(C) of 0.24 which is similar to what we have seen for duration.

The “Loudness map”

- Bradford and London: smaller contrast in loudness
- Leeds and Cambridge: larger contrast in loudness.
- Newcastle and Belfast.



We then applied MDS analysis to the confusion matrix to create the 'Loudness map' of our dialects.

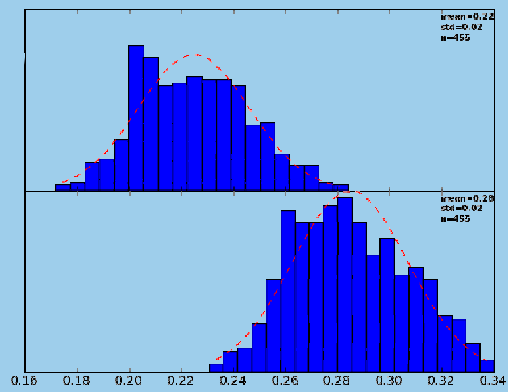
As you can see, we found yet another pattern of grouping. Noteworthy, Bradford and London, the two varieties with suspected 'syllable-timing‘, group together based on loudness and not on duration.

Remember though, that the overlap between different dialects is very large.

Loudness and duration

- The dialects show different grouping based on duration and loudness.
- Combining loudness and duration improves classification rate.
- There is still substantial overlap between dialects.

Duration only



Loudness and duration

Loukina & Kochanski
Durational variation in British dialects

13 September 2010
Page 45

Last, we combined loudness and duration to see whether together they would allow for better identification of dialects. The results are presented on this slide.

Conclusions

- English dialects can be identified above chance based on rhythmic properties, although the identification is far from 100%.
- There is greater overlap between dialects than between different languages.
- Bradford (Punjabi) speakers and London (Jamaica) speakers are usually distinguished from most other dialects in the corpus, but they **do not** show similar patterns of variation in duration.
- Belfast and Dublin English are consistently distinguished from each other.

The last two slides summarize our conclusions.

Conclusions

- Different groupings of dialects can be observed based on different properties.
- The grouping based on loudness differs from grouping based on duration or grouping based on f_0.
- There is no single ‘rhythm axis’: different grouping based on different properties means that there are no consistent ‘rhythm classes’ within British dialects.
- Rather then speaking about differences in ‘rhythm’ between dialects, we should speak about differences in particular properties such as loudness, duration of vowels or duration of consonants.

Acknowledgement

This project is supported by the Economic and Social Research Council (UK) via RES-062-23-1323.

Further information:
www.phon.ox.ac.uk/speech_rhythm

Thank you!