

## **Towards a typology and analytical framework for mixed-language texts**

**Mark Sebba**  
Lancaster University  
M.Sebba@lancaster.ac.uk



## **Code-switching in written texts: an under-researched area**

- Since 1970s, a huge boom in research on code-switching
- ‘Conversational code-switching’ is the main focus of almost all the published research

## **Code-switching in written texts**

- Some studies use written data as a replacement for conversational data, or as an additional source of data
- Fewer studies have written code-switching as the sole focus
- Written code-switching is often derided as ‘not authentic’ – spoken language is the ‘real thing’

## **Code-switching in written texts**

- Analytical frameworks used are usually ones developed for spoken code-switching
- Large amount of research remains unpublished: undergraduate dissertations, MA theses etc.

## **Recent developments**

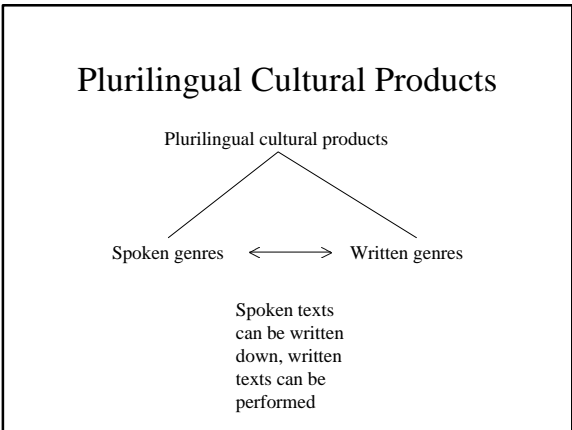
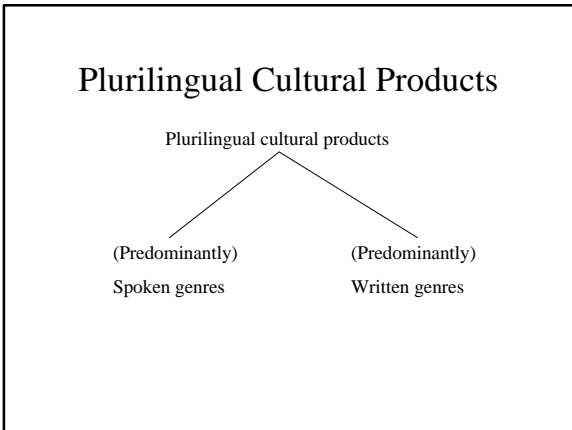
- New technology-based genres such as email have made new types of code-switching data available in large amounts
- Multilingual genres emerging as a result of globalisation, e.g. advertising
- Developments in corpus linguistics have made some new types of study possible

## **Nevertheless...**

Code-switching in written texts remains a relatively under-researched area.

- Unlike spoken code-switching, written code-switching has been resistant to a 'comprehensive' theoretical treatment
- This may have led to a relative lack of publications in this area and a feeling that it lacks 'respectability'

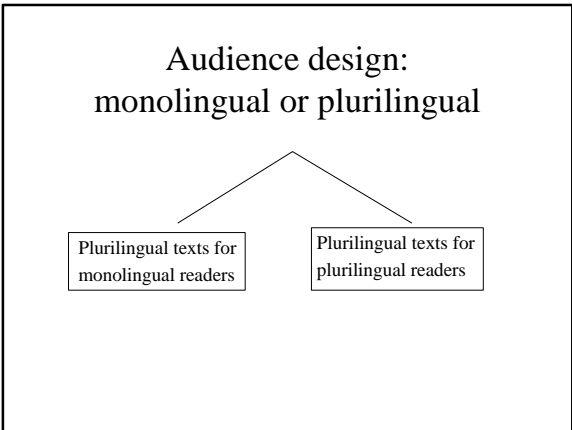
- ### In this paper I will
- try to situate the study of code-switching within a broader field which deals with the semiotics of 'plurilingual cultural products'
  - start to develop a typology which allows written code-switching to be analysed on a number of dimensions
  - discuss the relevance of various theoretical approaches which have been developed for spoken code-switching or other types of linguistic data



Jewish song from the Viciebsk gubernia with elements of Hebrew, Yiddish, and Russian:

Hebrew	Al Chet sechoton u l'fonecho B'eines u-v'Rocein.	על חטא שחטאנו לפיך באונס וברצון.
Russian	Prosti mnie, Gospodi, pomiluj Čto ja sogrešil Tebie S nuždoj I bies nuždoj!	פראָסט מיניע, גאָספּאָדי, פּאָמילוי, טשטש יא סאָגרעשיל טעביע ס'נוזשדזשי אי בעז נוזשדזשי!
Yiddish	Entfernt men ihm fun eiben, entfernt men ihm fun eiben Durak – durak! Kak tebe nie stidno, Kak tebe nie stramno Sogrešit -- Bies nuždoj!	ענטפערט מען אים פֿון איבן, ענטפערט מען אים פֿון איבן-- דוראַק – דוראַק! קאַק טעביע נייע סטידנאָ, קאַק טעביע נייע סטראַמנאָ סאָגרעשיט -- בעז נוזשדזשי!

Tamushanski, Robert J.: Belarusian-Yiddish Writings: Interesting Material for Ethnography and Philology  
ZAPISY Belarusian Institute of Arts and Sciences in the U.S.: <http://belarus8.tripod.com/litvaki/yiddish.htm>



Plurilingual texts for monolingual readers



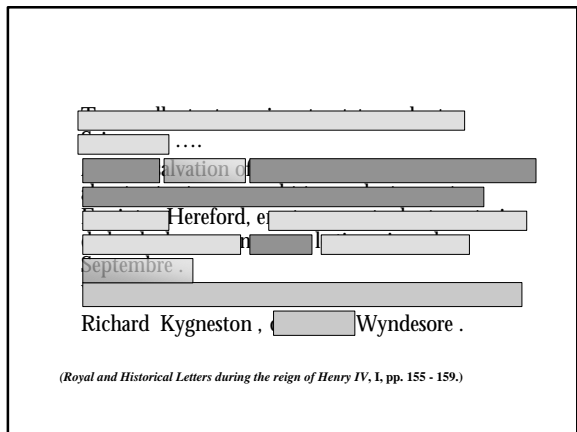
Plurilingual texts for monolingual readers



Extract from a Letter to Henry IV of England (1403)

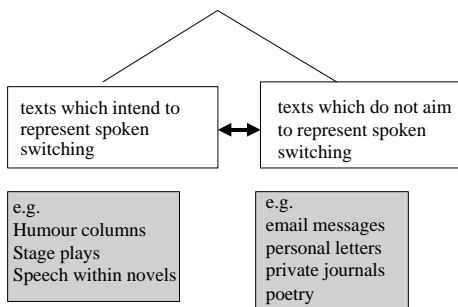
Tresexcellent , trespuissant , et tresredoute  
 Seignour , ....  
 And for salvation of youre schire and marches al  
 aboute , treste ye nought to no leutenaunt .  
 Escrip a Hereford, en tresgraunte haste, a trois  
 de la clocke apres noone , le tierce jour de  
 Septembre .  
 Vostre humble creatoure et continuelle oratour  
 Richard Kynneston , deane de Wyndesore .

*(Royal and Historical Letters during the reign of Henry IV, I, pp. 155 - 159.)*



*(Royal and Historical Letters during the reign of Henry IV, I, pp. 155 - 159.)*

Intention to represent speech



Intention to represent speech

- This is a separate issue from the question of whether a text is actually more 'speech-like' or 'writing-like'
- e.g. emails may be more 'speech-like' on various measures, speech in novels may have some characteristics of writing
- Biber found that no single dimension of co-occurring features distinguishes spoken and written texts.

## Intention to represent speech

- For some researchers only texts which represent spoken switching are interesting (and only if 'authentic')
- But these are only a subset of written texts which involve mixed languages
- Historically more genres allowed language mixing

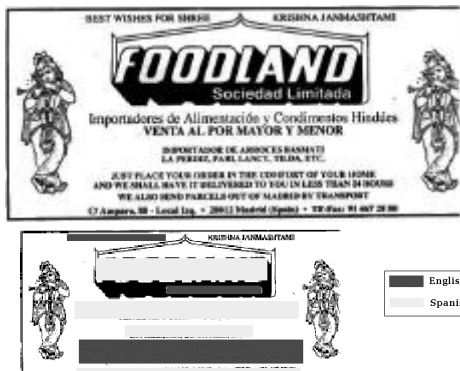
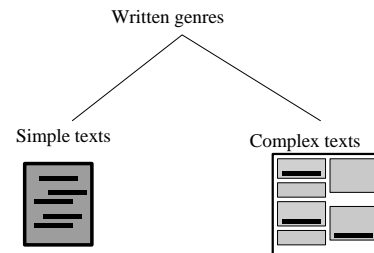
## Regulatedness of the genre



The most highly regulated space is at the centre  
Less regulated genres are in the outer circles

- 'Regulated spaces' – strictly monolingual norm within individual texts (e.g. newspaper articles)
- 'Unregulated spaces' – mixing of languages is 'allowed' within texts (e.g. personal emails)

## Simple vs. Complex Texts



## Relevance of layout

- Layout (e.g. fonts, font sizes, juxtaposition of images and position on the page) is highly relevant to the 'reading' of the codes in some genres
- Layout seems relatively unimportant in 'plain text' texts such as emails.

## Relevance of layout

- Even where layout is unimportant, *position* of elements within the text (e.g. beginning, end) is likely to be a significant analytical category



## Syntactic/Morphological models

- e.g. Matrix Frame model (Myers-Scotton)
- Written code-switching data tends to be used to supplement spoken data, or for comparison

## 'Markedness model' (Myers-Scotton)

- 'rights and obligations', 'rational choice'
- could be applied to more conversation-like interactive data, e.g. online chat
- harder to apply to other types of written data which are less interactive

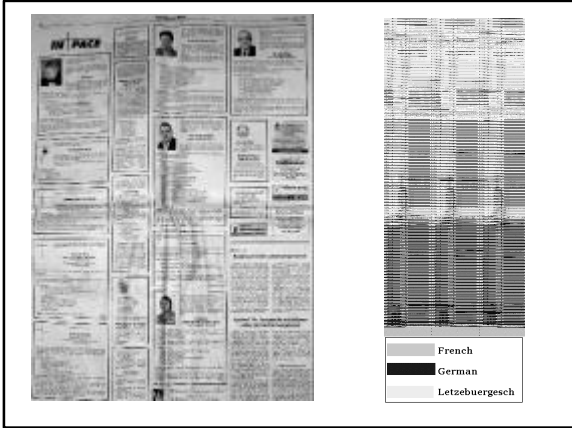
## Conversation analysis (Auer et al.)

- could be adapted to more conversation-like interactive data, e.g. online chat
- very difficult (impossible?) to apply to other types of written data which are less interactive

## *The interrelatedness of spoken language practices and literacy practices:*

- ***texts can contextualise and be contextualised by plurilingual speech and literacy practices***

Jones, Kathryn Eleri. (1999) Text, talk and discourse practices: exploring local experiences of globalisation. Ph.D. (Lancaster).



**The interrelatedness of spoken language practices and literacy practices:**

- *texts (mono- or plurilingual) can be mediated by plurilinguals*

Baynham, M. 1993 'Code switching and mode switching: community interpreters and mediators of literacy.' In: Street, B. (ed) 1993 *Cross-Cultural Approaches to Literacy*. Cambridge University Press. (pp. 294-314.)

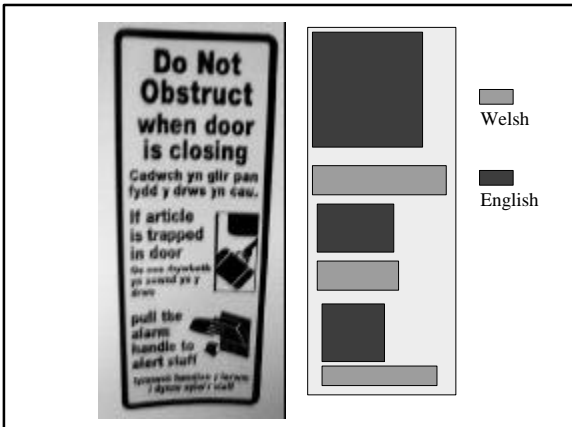
**Other possible approaches**

*Multimodal Semiotics: taking into account the visual images as well as the layout of text (and text as image):*  
 Kress and van Leeuwen  
*Reading images : the grammar of visual design* (1996)  
*Multimodal discourse : the modes and media of contemporary communication* (2001)

**Other possible approaches**

*Critical Discourse Analysis* : relationships between 'discursive practices, events and texts', and 'wider social and cultural structures, relations and processes'

- Emphasis on relations of power



**Conclusions**

- A 'multimodal' approach is needed to deal with many mixed-language texts
- Probably no existing framework on its own is sufficient
- The most useful approach may be an 'ethnographic' one which places each text in an overall social and cultural context before offering a 'micro' analysis