Dimensions of language variation world-wide

Josef Schmied
English Language & Linguistics
Chemnitz University of Technology
http://www.tu-chemnitz.de/phil/english/Schmied
Social constructionism and social constructivism are sociological theories of knowledge that consider how social phenomena or objects of consciousness develop in social contexts. Within constructionist thought, a social construction (social construct) is a concept or practice that is the construct (or artifact) of a particular group. When we say that something is socially constructed, we are focusing on its dependence on contingent variables of our social selves rather than any inherent quality that it possesses in itself. Thus our notion of a "cousin," what this terms includes and doesn't include and what it means to us does not exist "out there" in the world, but only in and through the social institutions that give it meaning within a culture. The underlying assumptions on which social constructivism is typically seen to be based are reality, knowledge, and learning.

Social constructs are generally understood to be the by-products of countless human choices rather than laws resulting from divine will or nature. This is not usually taken to imply a radical anti-determinism, however. Social constructionism is usually opposed to essentialism, which defines specific phenomena instead in terms of inherent and transhistorical essences independent of conscious beings that determine the categorical structure of reality.

A major focus of social constructionism is to uncover the ways in which individuals and groups participate in the construction of their perceived social reality. It involves looking at the ways social phenomena are created, institutionalized, known, and made into tradition by humans. The social construction of reality is an ongoing, dynamic process that is (and must be) reproduced by people acting on their interpretations and their knowledge of it.
0. Basics

Language systems are cognitively constructed; varieties, standards, accents, etc. are socially constructed.

How do we construct research?

What
- is useful?
- makes my life easier?
- allows me to say something intelligent that cannot be attacked easily?

Concepts - models - methodologies - theories

Research issue – question – topic/project – title
(+ operationalisation)
Research Paradigm

I = issue, i.e.
  new, i.e. not enough research yet
  relevant, i.e. important in discourse or for practical applications
  focussed, i.e. small, but data-based; careful and reliable

M = methodology, i.e. previous research, i.e. literature review incl.
  evaluation
  hypotheses possible?
  data base
  tests/procedure

A = analysis, i.e. examples as evidence
  statistical tables as summaries
  significance to generalise

C = conclusion, i.e. summary, e.g. in conclusion, we have shown
  interpretation, e.g. this proves that
  contextualisation, e.g. in a wider perspective
  outlook, e.g. further research is necessary
1. Typology of Englishes: concepts

ENL = English as a native language
ESL = English as a second language
EFL = English as a foreign language
EIL = English as an international language
ELF = English as a *lingua franca* (plural: *lingue franche*)
esp. of international science and technology
ESP = English for specific purposes
EAP = English for academic purposes
Why English? – “all languages are created equal” but ... English is more equal than others

*academic*:

- English has a more elaborate spectrum of world-wide functions, esp. as an additional language (EIL, ESP/EAP)
- English is more widely spoken than other languages
- English is the language of modern world communication technology
- English is related to German, but has undergone more (radical) changes in grammar, vocabulary, pronunciation, etc.
- English linguistics has more and better literature than other languages

*practical* (post-academic):

- English has more job offers in “language services”, teaching, translation (simultaneous interpretation), web publishing, information retrieval, etc.
- English skills (advanced and aware) are relevant in all jobs in an international intercultural context
2.2 Language variation according to users

2.2.1 Regional and social dialects

2.2.1.1 International varieties
British and American preferences, but also Canada, Australia/New Zealand (\textit{wellies} - \textit{gummies} [\textit{gumboots}] NZ)
e.g. in lexicon: \textit{luggage} - \textit{baggage}, \textit{lift} - \textit{elevator},
in grammar: \textit{have you got/eaten}; \textit{do you have/did you eat}?

2.2.1.2 Intranational varieties (dialects)
e.g. Geordie speech used in television programmes
(stereotyped notions in creative literature to create character and setting)
famous isoglosses (dialect lines) in pronunciation:
NEngland \textit{but, grass}; WEngland/New England: postvocalic \textit{<r>}
Cockney: systematic vowel shift (\textit{team} > \textit{tame} > \textit{time}),
  rhyming slang (\textit{trouble} and \textit{strive} = \textit{wife})
non-standard grammar less regional, more social (see 2.3.)

2.2.1.3 Cross-continental dialect chains show the arbitrariness of the terms
‘language’ vs. ‘dialect’ (North-South continuum)
e.g. Scottish English: \textit{I won't do it} - \textit{I'll not do it}
2.7 Language change

2.7.1 Speaker innovation and variation
spontaneous vs. network-specific
by imitation, innovations spread by adoption, diffusion in community
networks (weak ties of link person as innovation bridge -> 2.8.3)
e.g. which - witch, whether - weather
Northern Ireland: man [mo:n] - map [ma:p], mo'er
(cf. letters-to-the editor)

2.7.2 Social marking postvocalic [r]

2.7.3 Spread (wave metaphor)
style to style (casual > formal, incl. spoken > written)
word to word (lexical diffusion): e.g. today/advanced RP:
really = rarely, fear = fair

2.7.4 Reasons
status: particularly upper working class (less consciously) from neighbouring
communities with greater social status
sex: men as innovators tend to introduce vernacular, women prestige forms
(face-to-face) interactions
Major processes of change at work in/on modern English

1. Blurring of the class structure
2. Education for all (university expansion in 1990s)
3. Introduction of public broadcasting (provides a model for oral E = RP) (BBC “the wireless” 1922, TV 1953 after E2 Coronation)
4. Proliferation of film and video (Multimedia: 1927 film + sound, etc.; “Americanisation”)
5. Popular Youth Culture (“decline in formality”, “decline in standards”)
6. Decline of rural dialect and the rise in urbanisation (industrialisation, mobility, communication)
7. Global communication (new forms – new technical developments) (the internet: WWW, web 2.0: blogs, wikis, P2P/forum, etc.)

http://www.putlearningfirst.com/language/04change/process.html
5. Corpus Linguistics

5.1 Definition

\[ corpus_1 = \text{body or collection of written or spoken material upon which linguistic analysis is based} \quad (\text{cf. Bloomfield structuralism}) \]

\[ corpus_2 = \text{machine-readable} \]

\[ \quad \text{‘representative’, i.e. stratified ‘model’} \]

\[ \quad \text{i.e. more than a text collection!} \]

for computer-based language analysis:

\[ \text{corpus-informed (lang. awareness/ELT)} \]

\[ \text{vs. corpus-driven (research)} \]

with tools/corpus-analysis software:

\[ \text{WordSmith (with ICAME CD)} \]

\[ \text{Sara (with BNC in TUC Bib)} \]

\[ \text{AntConc:} \]

http://www.antlab.sci.waseda.ac.jp/antconc_index.html
5.2 Reasons

for the popularity of corpus linguistics esp. among non-native speakers because it combines a qualitative and quantitative perspective

- citations used as samples of language
- provides a view beyond individual experience
- rules out individual salience
- computer processable

output:
- concordances (KWIC = key word in context)
- collocates
- relative vs. absolute frequencies
Developments in corpus compilation:

reference data

1950s American structuralists, e.g. Harris

1959 Quirk: Survey of English Usage (SEU)

  1,000,000 words written/spoken 1953-1987

  > London-Lund corpus of spoken English

1963/64 Francis/Kucera: Brown Corpus

  1M of written American English from 1961


  1M written BritE (Lancaster-Oslo/Bergen Corpus) from 1961

1980 - Cobuild Corpus (Birmingham, Sinclair) → Bank of English

1990 - International Corpus of English (ICE):

  UK, US, EA (KE/TZ), ZA, HK, AU, NZ, PH

1990 - International Corpus of Learner English (ICLE)

1990 - 1993 British National Corpus 100M (10M spoken)

5.4 Corpus compilation principles (cf. user manuals!)

large and stratified (balanced):

- mega-/reference corpora
  - British National Corpus (BNC)
    90 Mio written/10 Mio spoken, demographic/context-governed from 1991-94
    http://www.natcorp.ox.ac.uk/
  - American / Australian National Corpus being compiled now
- ‘national corpora’, e.g. ICE http://www.ucl.ac.uk/english-usage/ice/
  e.g. ICE-East Africa: http://ell.phil.tu-chemnitz.de/ice-search/
- genre/domain specific corpora
  e.g. SPACE (Specialised & Popular ACademic English),
  Trains (dialogue corpus)
- translation corpora
  e.g. EU corpus: http://ell.phil.tu-chemnitz.de/multiSearch/
  English-German Translation Corpus http://ell.phil.tu-chemnitz.de/search/

‘quick and dirty’/ad-hoc corpora, e.g. for translation problems
5.5 Corpus search strategies

Investigating Context
• concordance
• collocation = the appearance of one particular word form in certain distance of another particular word forms > different meanings can have different collocates
• colligation = the appearance of one particular word form in a particular grammatical structure (e.g. *cases*)
• connotation = the semantic environment, can have positive or negative value (‗semantic prosody‘)
  e.g. *happen, cause, attempt, try, fail*

Exploring Vocabulary
• absolute vs. relative frequency of a word form
• (standard) deviation from mean frequency of word forms
5.6 Corpus research examples

- How frequent is a particular morphological form/grammatical structure?
- Which particular structures have particular meanings?
- Which particular structures have particular locations in texts?

Corpus tasks have degrees of complexity

- relevance of tagging:
  - parts-of-speech (POS), e.g. CLAWS tagging for LOB
    ([http://ucrel.lancs.ac.uk/claws/trial.html](http://ucrel.lancs.ac.uk/claws/trial.html))
  - semantic → semantic web/web 3.0
Language analysis: tagging and annotation

part-of-speech tagging from LOB (CLAWS1 tagset):

- hospitality_NN is_BEZ an_AT excellent_JJ virtue_NN , , but_CC not_XNOT when_WRB the_ATI guests_NNS have_HV to_TO sleep_VB in_IN rows_NNS in_IN the_ATI cellar_NN ! ! the_ATI lovers_NNS , , whose_WP$ chief_JJB scene_NN was_BEDZ cut_VBN at_IN the_ATI last_AP moment_NN , , had_HVD comparatively_RB little_AP to_TO sing_VB ' ' he_PP3A stole_VBD my_PP$ wallet_NN ! ! ' ' roared_VBD Rollinson_NP . .

syntactic annotation/treebanks

- square brackets (sequential/horizontal, no indentation; Lancaster):

from: ttp://www.comp.lancs.ac.uk/computing/research/ucrel/annotation.html#treebank
8. Language & Culture

8.0 Definitions

culture = transgenerational practices through which humans communicate

culture = a network of habitual disparities

culture = a system of symbols by which humans enact embodied understanding (Geertz); cultural meanings are not individual but shared, public

Geertz 1973 proposes cultural relativism, but does not deny universals
if cultural universals exist, they must be abstract, insubstantial

culture = a system of knowledge, what one has to know in order to act acceptably (Goodenough 1996)

Raymond Williams (1958) moves “Cultural studies” from “High Culture” to “Ordinary Culture”
8.1 Locating Culture on different levels of language

- pronunciation → complex (vowels/consonants, intonation)
- lexicon → colours, esp. green (spectrum and metaphors)
- syntax / grammar →
  - conceptualisations (mass/count)
  - modality to express politeness (could you do this for me?)
  - “verititative” to express that s.th. is true (in English, part of “modality”)
- text / discourse →
  - hedges in academic writing (I think, they claim)
  - deixis in demonstrative systems (this, that, archaic: yonder)
  - argumentation structure:

  - Kaplan (1966) paragraphing strategies reflect cultural styles (diagram)
  - Galtung (1985) hypothesized 4 different “intellectual styles”: Saxonic, Gallic, Nipponic, Teutonic
linguistic relativity n. The hypothesis first suggested by the German ethnologist Wilhelm von Humboldt (1767–1835) and reformulated in the 20th century by the US fire prevention officer and linguist Benjamin Lee Whorf (1897–1943) that people who speak different languages perceive reality and think differently, because categories and distinctions encoded in one language are not necessarily available in another.

E.g., the language of the North American Hopi Indians has no forms corresponding to English tenses and no words for conceptualizing time as a dimension; therefore, Whorf argued, it would be difficult for a Hopi and an English-speaking physicist to understand each other's thinking about time.

See also linguistic determinism, Sapir—Whorf hypothesis from www.encyclopedia.com
8.3 Culturally determined patterns: scripts and frames

psychological constructs as atomic forms of human **stereotyped** knowledge

frames
- are knowledge structures that contain fixed structural information
- have slots to accept a wide range of values
- trigger comprehension and recall

scripts
- are formalized representations of complex actions
- "consists of concepts and relations between concepts" (Schank)
- list of primitives (stereotype procedures) e.g.

```
frame: HOUSE
subframe of: building
Is-part of: village, city, suburb
material: wood, stone, concrete
# of windows: integer, >2
# of doors: integer, default 1
```
8.4 Cross-cultural communication

8.4.1 Introduction

**definition intercultural**

Although Wikipedia 2009 says: “Intercultural communication principles guide the process of exchanging meaningful and unambiguous information across cultural boundaries” the term is ambiguous, because it may be the same as cross-cultural (s. below)

OR it may emphasize that in intercultural communication a new communication context is created that constitutes more than a compromise between the two (or more) cultures of its communication participants

**definition cross-cultural**

= how people from different cultural backgrounds endeavour to communicate effectively
danger: misinterpretations of social meaning can result from unrecognised differences

**cross-cultural pragmatics**

in the Japanese context in response to a kindness, instead of the speech act of thanking, the speech act conventionally performed is that of excusing - excusing for the trouble that one has caused the other person.

requests in German tend to be made syntactically in a more direct manner than in English (House / Kasper 1981).
8.4.2 10 simple tips for cross cultural communication

Slow Down = Even when English is the common language in a cross cultural situation, this does not mean you should speak at normal speed.

Separate Questions = Try not to ask double questions such as, “Do you want to carry on or shall we stop here?”

Avoid Negative Questions = Many cross cultural communication misunderstandings have been caused by the use of negative questions and answers.
   In English we answer ‘yes’ if the answer is affirmative and ‘no’ if it is negative.
   In other cultures a ‘yes’ or ‘no’ may only be indicating whether the questioner is right or wrong. For example, the response to “Are you not coming?” may be ‘yes’, meaning ‘Yes, I am not coming.’

Take Turns = Cross cultural communication is enhanced through taking turns to talk, making a point and then listening to the response.

Write it Down = If you are unsure whether something has been understood write it down and check.

Effective cross cultural communication is about being comfortable. Giving encouragement to those with weak English gives them confidence, support and a trust in you.

Check Meanings = When communicating across cultures never assume the other party has understood.
   Be an active listener. Summarise what has been said in order to verify it. This is a very effective way of ensuring accurate cross cultural communication has taken place.

Avoid Slang

Watch the humour = When using humour think whether it will be understood in the other culture.

Maintain Etiquette = cross cultural awareness training or at least some research on the target culture.

http://www.kwintessential.co.uk
**8.4.3 Text-type-specific examples from EAP**

**email:**

“Dad; I must say that I am so happy to read from you. The mail is so inspiring. I must thank you immensely for the love you have for me. From the outset, I knew that your greatest desire was to see me an accomplished scholar. The decision to bring me on board this project, testifies to this. A Big thank you for this. I will send samples of the data to as soon as I finish. I think I can confidently handle the task.”

**scholarly article:**

<B1/>Bilingualism is an ability of individual of using more than one language correctly (T). <B2/>In this situation individual could use more than one language interchangeably and appropriately (N+). <B3/>This means one could use language according to the context (N+?). <B4/>In bilingualism there are four [1-4?] things which should be <_beared>_ <+_borne> in mind and this is ...
为了避免意外落水，
请家长看护好自己的小孩，不要玩水。

In order to avoid the accident falling in the water, asks the guardian to nurse own small child, do not want to play the water.

为了避免意外失足，
请家长看护好自己的小孩。

In order to avoid the accident losing footing, asks the guardian to nurse own child.
Kachru: 3 circle model of World Englishes

**The Inner Circle**
Countries: UK, USA, Canada, Australia, New Zealand
-> ENL countries

norm providing
Kachru: 3 circle model of World Englishes

**The Outer Circle**
Countries: Bangladesh, Singapore, India, Nigeria, Malaysia, ...

-> ESL countries

norm developing
The Expanding Circle
Countries: China, Egypt, Indonesia, Taiwan, Korea, Israel,...
no official status
-> EIL countries

norm dependent
Kachru: 3 circle model of World Englishes

**Criticism**

Model implies uniformity of countries within one circle

-> not true: even in Inner Circle differences in linguistic diversity

today’s migration left out of account

gradients between Outer and Expanding Circle -> transition Argentina, Belgium, Denmark?
McArthur: circle of World English

**Inner Circle:**
- World Standard English
  -> but not in identifiable form

**Outer Circle:**
- Band of regional varieties of English