Measuring the Borrowability of Word Meanings

Uri Tadmor and Martin Haspelmath

Max Planck Institute for Evolutionary Anthropology
(Jakarta and Leipzig)

Frequently made assertions

○ "Nouns are more often borrowed than verbs"

○ "Basic vocabulary is rarely borrowed"

○ "Content words are more often borrowed than function words"

○ "Terms for body parts are difficult to borrow"

○ "A language cannot borrow into a closed set"

But so far no large cross-linguistic study has been undertaken to examine these assertions.

The Loanword Typology Project

○ A collaborative project coordinated by the Department of Linguistics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

○ Headed by Martin Haspelmath and Uri Tadmor

○ Aimed at studying lexical borrowing patterns in a representative sample of languages from around the world
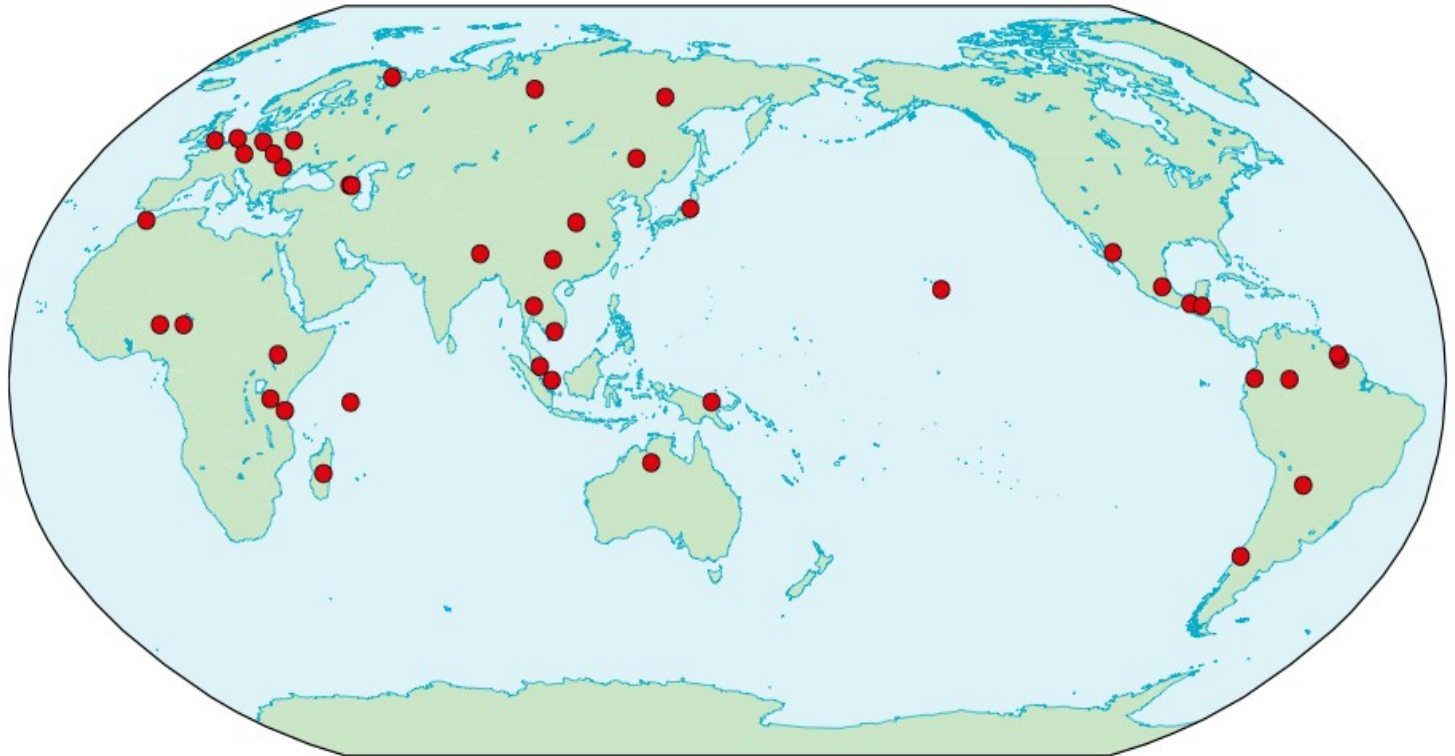
The Loanword Typology Project

The LWT project has two planned results:

○ An electronic database of loanwords in 40 languages, based on a meaning list of 1464 lexical meanings (to be published online)

○ A book with 40 case studies and general studies of lexical borrowing

Languages

40 languages from different parts of the world

The contributors

Contributors:

- Contributors must be specialists of the language and their history

- Contributors had to be willing invest the considerable amount of time and effort needed

- They were asked to provide counterparts of 1464 meanings, plus loanword status and additional information

- The project setup was discussed thoroughly at several workshops in Leipzig

The database

Authors' template:

The book

Each case study contains at least the following sections:

- The language and its speakers
- Sources of data
- Contact situations
- Numbers and kinds of loanwords
- Integration of loanwords
- Grammatical borrowing
- List of loanwords from the database

Preliminary results

Borrowability by ontological category

Nominal meanings	31.4%
Adjectival meanings	15.8%
Verbal meanings	14.4%
Adverbial meanings	8.5%
All lexical meanings	**25.4%**
Grammatical meanings	11.8%
All meanings	**24.5%**

Preliminary results

Semantic fields: high borrowability

Clothing and grooming	39.5%
Religion and belief	39.3%
The house	36.5%
Law	34.1%
Agriculture and vegetation	31.0%
Social and political relations	30.8%
Food and drink	30.3%

Preliminary results

Semantic fields: mid-range borrowability

Warfare and hunting	28.3%
Possession	27.2%
Animals	26.4%
Cognition	24.3%
Basic actions and technology	24.1%
Time	22.9%
Speech and language	22.5%
Quantity	20.5%
The physical world	20.3%

Preliminary results

Semantic fields: low borrowability

Emotions and values	19.8%
Motion	17.1%
Kinship	15.4%
Spatial relations	14.5%
The body	14.2%
Sense perception	11.6%

The 10 least borrowed meanings on the LWT list (all counterparts)

Meaning	Representation rate (no. of languages)	No. of word forms in db	% of word forms not borrowed
this	38	39.3	100%
up	38	36.5	100%
3SG pron.	35	28.5	100%
where?	39	43.5	99%
to lose	37	38.8	99%
there	38	51.0	98%
that	37	50.7	98%
here	39	49.8	98%
why?	39	46.3	98%
which?	39	41.8	98%

The 10 least borrowed meanings on the LWT list (unanalyzable counterparts only)

Meaning	Representation rate (no. of languages)	No. of word forms in db	% of word forms not borrowed
this	29	28.8	100%
to stand	29	27.6	100%
3SG pron.	33	25.5	100%
to rise	27	22.1	100%
to lie down	25	26.7	98%
to run	28	26.5	98%
that	27	33.0	97%
the nose	35	32.8	97%
to go	34	32.2	97%
yesterday	32	32.0	97%

The 10 most borrowed meanings on the LWT list (unanalyzable counterparts only)

Meaning	Representation rate (no. of languages)	No. of word forms in db	% of word forms borrowed
the school	27	30.3	100%
the bank	25	25.0	96%
the bus	31	33.6	94%
the machine	26	21.3	94%
the coffee	35	36.5	93%
the motor	31	30.5	93%
the television	26	27.5	93%
the tea	32	35.0	91%
the pen	27	33.0	91%
the towel	27	28.0	91%

Some generalizations

- 'Nouns' are more often borrowed than 'verbs'. In fact, 'nouns' > 'adjectives' > 'verbs' > 'adverbs'

- Lexical meanings are more often borrowed than grammatical meanings

- Among the least borrowable are demonstratives, personal pronouns, interrogatives, body parts, basic polysemous verbs

- However, anything can be borrowed, even into closed sets and highly structured systems like pronouns and (lower) numerals

Loanword Typology: The Future

- Placing the database online (early 2009)
- Publishing the book (late 2009)
- Continued contributions to online database
- Periodical revisions ('editions') of online database

THANK YOU