Typology and language contact: the case of verb borrowing

1. **What is typology?**
2. **Why do we need typology in contact linguistics?**
3. **Borrowing scales: how to read them, how to explain them**
4. **Verb borrowing in typological perspective**
5. **Evaluating predictions made by typological generalizations: the case of verb borrowing in Coptic**

Assumption:
The goal of linguistics is not only to describe linguistic phenomena, but also to **explain** them.

1. What is typology?

Typology is the study of the limits on linguistic diversity. If all structural possibilities of language were equally likely, then we would expect all possibilities to occur with roughly equal frequency. However, this is not the case. Some possibilities never occur, some occur infrequently, and some occur frequently.

Example: of the six **logically possible** basic orders of SUBJECT OBJECT VERB, all are attested. However, the relative frequency in a sample is as follows:

TABLE 1: DOMINANT WORD-ORDER FREQUENCIES IN A SAMPLE OF 1228 LANGUAGES (DRYER 2008)

DOMINANT ORDER	NUMBER OF LANGUAGES	FREQUENCY
SOV	497	40.5%
SVO	436	35.5%
VSO	85	0.07&
VOS	26	0.02%
OVS	9	0.01%
OSV	4	< 0.01%
NO DOMINANT ORDER	171	13.9%
TOTAL	1228	

Some possibilities are independent of other variables, while others are dependent.

Example, Greenberg's Universal 3:
Languages with dominant VSO order are always prepositional (Greenberg 1966: 78).

Typological classification represents the observation of an empirical phenomenon (language) and the classification of what we observe. Typological generalization – language universals – is the formation of generalizations over our observations. And the functional-typological approach constructs explanations of the generalizations over what we have observed. In this sense, typology represents an **empirical scientific** approach to the study of language (Croft 2003: 2).

Heavily **deductive** approaches, in which a generalization is proposed on the basis of limited data or due to the needs of a theoretical framework, are less useful, for several reasons.

איתן גרוסמן
לשונות במגע
13.4.2010

- One, sometimes the needed empirical evaluation is never carried out, so false generalizations are allowed to stand.
- Two, false generalizations often become entrenched in linguistic discourse, so linguists spend a lot of effort explaining why they are in fact not false.
- Three, they tend to obscure factors that might be unexpectedly important.
- Theoretical frameworks are subject to fashion, and what is commonly accepted wisdom today might look outdated if not ridiculous tomorrow. Empirical data should not be contingent on such transient objects.

Explanations can be 'internal' or 'formal,' i.e., explain generalizations due to the nature of linguistic competence, or can be 'external' or 'functional,' i.e., explain generalizations due to constraints on performance (e.g., pragmatics, general properties of cognition). Typologists tend to prefer 'external' or 'functional' explanations (Comrie 1984, Hyman 1984, Haspelmath 2002). This is largely due to the research paradigm established by Greenberg (e.g., Greenberg 1966).

Some largescale online databases
- WALS (World Atlas of Linguistic Structure), available online at http://wals.info/index
- WOLD (World Loanword Database), available online at http://wold.livingsources.org/
- The Universals Archive, available online at http://typo.uni-konstanz.de/archive/intro/index.php.

Some useful introductions to typology
Greenberg (1966), Ramat (1987), Croft (2003), Shopen (2007)

Some typologically-oriented studies of borrowing
Matras & Sakel (2007), Matras (2009), Haspelmath & Tadmor (2009)

2. Why do we need typology in contact linguistics?
In many works on borrowing, one finds absolute or relative statements about **borrowability**, what can be borrowed or how likely something is to be borrowed.

(i)   'Inflectional morphemes are not borrowed.'
(i')   'Inflectional morphemes are less easily borrowed than derivational morphemes.'
(ii)   'Verbs are never borrowed as verbs.'
(ii')   'Verbs are less easily borrowed than nouns.'

Problems:
These statements are usually (1) made in order to satisfy the requirements of a theoretical framework, (2) based on **anecdotal** evidence or **casual observation**, i.e., a language or group of languages familiar to the linguist, or (3) both. In such cases, there is little empirical basis for such strong statements, even if relativized ('more easily borrowed,' 'less frequent').

Moreover, these statements privilege linguistic structure over linguistic performance. As such, they will predict that the structural status of an element is especially relevant for its

איתן גרוסמן
לשונות במגע
13.4.2010

borrowability. This has been falsified: it has been demonstrated time and again that purely structural information makes bad predictions. Borrowing is highly dependent on pragmatic and sociolinguistic factors.

Another serious problem: linguistic categories, such as Dative or Verb, are not universal. Each language has its own categories, which linguists describe. Typologists have ways to deal with this problem, e.g., the use of **comparative concepts** (Haspelmath 2010). Comparative concepts are constructed for the purpose of cross-linguistic comparison, and make reference only to universal concepts, which may be formal (e.g., *X precedes Y*) or functional (*encodes a spatial relation*). It has recently been argued that comparative concepts are necessary for contact linguistics, since categories (e.g., adjective) are language specific (Grossman 2010b). Furthermore, the fact that borrowings often differ from **both** the 'same' category in the donor language and in the recipient language ('partial integration') is an additional demonstration that comparative concepts are necessary. Finally, bilingual speakers and translators make use of comparative concepts, since they can equate constructions in both languages that might be structurally very different.

Still, we can treat claims about borrowability as hypotheses that generate predictions that can be evaluated empirically. In linguistics, we are looking for **hypotheses** whose **predictions** can be evaluated **empirically** in order to reach **generalizations** about linguistic behavior. Furthermore, we would like to **explain** these generalizations. Explanations are stronger if they are **non-circular**. This means that they do not refer to the nature of language itself, but rather to language-external factors. Explanations can generate further predictions.

Example: adposition borrowing (Grossman 2010b)
Generalizations to be explained
(1) Adpositions seem to be resistant to borrowing, and they are borrowed only in languages that have borrowed nouns, adjectives, discourse markers, coordinating and subordinating conjunctions.

Ex. 1. Greek *katá* 'according to' > Coptic   (Grossman 2010b)
e-p-rôme	*ero-f*
to-the-man	to-3SG.M

kata-p-rôme	*kataro-f*
according_to-the-man	according_to-3SG.M

Ex. 2. Sicilian *favuri* 'in favor of' > Maltese *favur(i)*   (Saari 2002)
fuq-kom	*favuri-kom*
on-2PL	in_favor_of-2PL
'on you (pl.)'	'in your (pl.) favor'

Ex. 3. Kurdish *gal* 'with' > Jewish Neo-Aramaic (Arbel) *gal/gall-*   (Khan 1999)
gal-son-èu	*gall-eu*
with-grandfather-3SG.M	with-3SG.M
'with his grandfather.'	'with him.'

Ex. 4.   Hebrew *lixvod* > Yiddish *lekoved*

lixvod ha-šana	*ha-xadaš-a*		*lekoved*	*der*	*nayem yor*
in_honor_of:CS the-year:SG.F	the-new-SG.F		in_honor_of	the.DAT	new.DAT year
'in honor of the new year.'			'in honor of the new year.'		

Possible explanations (i marks a formal explanation, ii a functional one)

(i)   Adpositions are a grammatical (i.e., non-lexical) category. Grammatical categories are harder to borrow than lexical categories.

(ii)   Adpositions, especially basic spatial ones like 'in,' 'on,' or 'to,' are frequent in discourse. As such, they are 'protected' by routine and not likely to be affected by borrowing.

Predictions:

(i)   The more grammatical a category is, the harder it is to borrow.

(ii)   The more frequent an adposition is, the less likely it is to be replaced by borrowing. Since basic spatial adpositions are especially common in discourse, they will be least affected by borrowing, while other non-basic spatial ('between,' 'alongside') or non-spatial ('according to,' 'without') meanings will be more likely to be replaced by borrowing.

(2) Adpositions seem to be borrowed only in cases of intensive language contact with considerable or universal bilingualism. N.B. This observation is relevant to **spoken** language contact. In **written** language contact and contact via **translation**, the situation can be much more complex.

Possible explanations

(i)   ?

(ii)   Basic spatial adpositions are unlikely to be associated with contexts that are specific to the donor language. Other adpositions might be more closely associated with culturally or linguistically specific communicative modes, e.g., ('despite' in argumentative discourse, which might be characteristic of a donor language that is dominant in academic and official discourse).

Predictions

(i)   ?

(ii)   The type of discourse influences whether (and which) adpositions are borrowed.

3. Borrowability scales

'An important question that has often been raised in discussions of borrowing is whether there are inherent differences in the likelihood of different word classes, categories, or types of morphemes to be affected by borrowing' (Matras 2009: 153).

איתן גרוסמן
לשונות במגע
13.4.2010

<u>Muysken (2008: 177) partly based on Haugen (1950).</u>
nouns > adjectives > **verbs** > coordinating conjunctions > adpositions > quantifiers > determiners > free pronouns > clitic pronouns > subordinating conjunctions
<u>Matras (2007, 2009)</u>
nouns, conjunctions > **verbs** > discourse markers > adjectives > interjections > adverbs > other particles, adpositions > numerals > pronouns > derivational affixes > inflectional affixes

How are borrowability scales usually read?
- **Quantitative reading**: a language will borrow **more** of elements to the left than elements to the right.
- **Temporal reading**: a language will borrow elements to the left **earlier** than elements to the right.
- **Implicational reading:** if a language has borrowed an element, it **implies** (or **entails**) that it has also borrowed all elements to its left.
- **Frequency reading:** borrowing of elements to the left is **more common** cross-linguistically than borrowing of elements to the right.

But:
- Borrowing scales have to be established empirically, i.e., on the basis of a large set of cross-linguistically representative data.
- They do not constitute explanations in themselves, but need to be explained.

'While such hierarchies provide a rough indication of possible predictions on borrowing, their meaningfulness is debatable as far as an explanatory account of the borrowability of categories is concerned. On a word-class hierarchy, there is not always an obvious link between any two categories that occupy any two positions, whether they are adjacent or not' (Matras 2009: 157-158).

What could constitute an explanation?
- Universal grammar: the most frequently/easily borrowed elements are also the most universal, i.e., nouns are the most universal lexical category (Muysken).
- Intensity of contact: intensity of contact (e.g., prestige) leads to more borrowing (Thomason & Kaufman).

Casual contact	Category 1	content words
	Category 2	function words, minor phonological features, lexical semantic features
↓	Category 3	adpositions, derivational suffixes, phonemes
	Category 4	word order, distinctive features in phonology, inflectional morphology
Intense contact	Category 5	significant typological disruption, phonetic changes

- Increased bilingualism leads to more contexts in which all of the bilingual speaker's repertoire can be exploited for communicative purposes (Matras).

איתן גרוסמן
לשונות במגע
13.4.2010

4. Verb borrowing

[a] lexical item whose meaning is verbal can never be included in the set of borrowed properties (Moravcsik 1978: 111). [= verbs cannot be borrowed as such, but must be borrowed as nouns and 'reverbalized' in the borrowing language, Moravcsik 2003, cited in Wohlgemuth 2009]

In the absence of a comprehensive survey or even a systematic sample of languages that borrow verbs, it is difficult to either confirm or refute the existence of a [borrowing] hierarchy (Matras 2009: 176).

Recent typological studies
- Wichmann & Wohlgemuth 2008, Wohlgemuth 2009

The present work is meant as a contribution to the research on verb borrowability, investigating and discussing the topics raised by the following key questions:
1. Why do many languages seem to have more difficulties borrowing verbs than nouns?
2. Due to which factors do languages apparently borrow more nouns than verbs?
3. Can verbs be borrowed *as verbs* or must they (always) be "re-verbalized" in the borrowing language?
4. By which mechanisms and paths are verbs being borrowed and, if necessary, adapted?
5. Is the choice of these mechanisms dependent on linguistic and/or extralinguistic factors in the donor and/or the recipient languages?
6. Which factors are these and what effects do they have in individual languages as well as cross-linguistically?
7. Are there universal constraints on verb borrowability?   (Wohlgemuth 2009: 4-5).

Answers given:
1. These difficulties are more apparent than real. Almost all languages in the sample borrow verbs in some way. In some languages, especially those with gender systems, nouns involve greater difficulties than verbs in terms of integration.
2. A number of reasons are offered, some cognitive and semantic.
3. Verbs can definitely be borrowed as verbs, in a number of different ways. The 'reverbalization' hypothesis is indefensible on empirical grounds.
4. See below.
5. Considerations like 'morphological compatibility' have been overestimated in the literture. However, a surprising result turned up. The different strategies tend to correlate with basic constituent orders: head-dependent orders tend to correlate with direct insertion, while dependent-head orders with light verbs. Extralinguistic (sociolinguistic) and functional (discourse) considerations play a large role.
6. Contact situation, universality of bilingualism, language attitudes, etc. See Wohlgemuth (2009) for a full discussion.
7. Not really.

Sample: 'all you can get' rather than 'convenience' or 'balanced'
- 533 language pairs (donor > recipient)
- 140 donors
- 352 recipients
- 429 total languages

Methods of data collection (Wohlgemuth 2009: 37)
– "classic" literature research in various libraries
– internet research using Google Scholar, Google Books, OVID etc.
– requests sent to the *Linguist List* and the replies thereto
– personal discussions following presentations on my research at conferences

איתן גרוסמן
לשונות במגע
13.4.2010

– personal communication with colleagues working on particular languages – data contributed to the Loanword Typology Project by colleagues – data submitted by colleagues using a dedicated web submission form.
– serendipity; e.g. overhearing speakers using loan verbs or ad hoc/nonce forms, or coming across loan verb examples in scholarly papers on an entirely unrelated topic

1. Verb borrowing strategies (for detailed discussions of each strategy, see Wohlgemuth 2009).
2. How to interpret situations in which there are multiple strategies that coexist synchronically.

'Input forms' (donor language) - cut across borrowing strategies
- Abstract form: bears no inflection or is not a regular form of the verb.
- General form: a form that lacks inflection and is similar to or identical with a bare stem.
- Infinitive and verbal nouns
- Participles
- Imperative (?)
- Inflected forms: 3rd person sg

Verb borrowing strategies (Wohlgemuth 2009: 293)
i)   Light verb strategy
…. a borrowed verb is to enter it as an non-inflecting part into a complex predicate, joining a native verb which takes all the inflection.

Ex. 5   BOHAIRIC COPTIC < GREEK   (Grossman 2010a)
a-f-er-klêronomin
PAST.AFF-3SG.M-do-inherit\INF^greek
'He inherited….' (Gr. *klêronomein*)

Ex. 6   TURKISH < FRENCH   (Lewis 1967: 154)
isole etmek
isolate to_do/make
'to isolate, insulate' (Fr. *isoler*)

Ex. 7   YIDDISH < HEBREW   (Matras 2009)
máskim zayn
agree to_be
'to agree'

Ex. 8   DOMARI < ARABIC   (Matras 2009: 181)
fhimm-ho-mi
understand-become-1SG
'I understand'

(ii) Indirect insertion

… a verbalizer of some kind is applied so that the verb form can be inflected.

Ex. 9	HUNGARIAN < GERMAN		(Moravcsik 1975)
	leiszt-ol	*ful-el*	
	accomplish-VERBALIZER	ear-VERBALIZER	
	'accomplish' (G. *leist-en*)	'listen hard'	

(iii) Direct insertion

the borrowed verb stem is simply used like a native one without any morphosyntactic adaptation.

Ex. 10	FRENCH < DUTCH	(Wohlgemuth 2009: 88)
	hiss-er	
	hoist-INF	
	'to hoise' (Dutch *hiss-en*)	

Ex. 11	QUECHUA < SPANISH	(Matras 2009: 177)
	balura-ni	
	value-1SG	
	'I value' (Sp. *valora-r*)	

Ex. 12	TASWAQ < TUAREG	(Matras 2009: 177)
	yízmàm	
	'to press' (Tuareg *y-əẓmăm* '3SG.PRFV-press'	

(iv) Paradigm transfer: verbal inflection is borrowed

the borrowed verb's inflectional morphology of the donor language is borrowed along with it, introducing a new inflectional paradigm into the recipient language.

Ex. 13	ROMANI (AGIA VARVARA) < TURKISH	(Bakker 2005: 9)
	and o sxoljo ka siklos te okursun ta te jazarsun	
	in ART school FUT learn.2 COMP read.2SG and COMP write.2SG	
	'at school you will learn how to read and write'	

-sun, the Turkish 2nd person marker, is borrowed along with the borrowed verbs. Turkish infinitives *okumak* 'to read', *yazmak* 'to write.'

General tendencies of strategy distribution (Wohlgemuth 2009: 150-151)
- These limitations aside, the following general tendencies or preferences with regard to the accommodation strategies and patterns applied can be identified rather safely.
- **Direct Insertion** is by far the most frequently used strategy when one counts the absolute number of languages applying it. It is also the strategy that is most frequently used in combinations and has accordingly the lowest percentage of languages exclusively using it.
- **Indirect Insertion**, ranking third by the number of languages using it, is clearly less frequently found combined with other strategies: close to half of the languages applying it use it as their sole

strategy.

- **The Light Verb Strategy** is somewhere in between the two other major strategies. It ranks second in absolute frequency and appears to co-occur with other strategies more frequently than Indirect Insertion.
- **Paradigm Insertion** is an extremely rare accommodation technique. Generalizations about the combined use of Paradigm Insertion and other strategies were avoided, because they would be based on too few data.

Interestingly, *all* combinations of more than two strategies involve Direct Insertion. Furthermore, of the ten languages using more than two strategies, nine actually use Direct Insertion and Indirect Insertion plus one (or two) other. The dyad *Direct Insertion + Indirect Insertion* is thus also the most frequent one, albeit not with a statistically significant margin.

If the attested use of multiple strategies is indicative of diachronic development, one might expect that Direct Insertion, as a less complex strategy, would have occurred later in the borrowing history of a recipient language than other strategies.

Borrowing of borrowing patterns

Languages can borrow a borrowing mechanism itself. For example, the Uto-Aztecan language Nahuatl uses *-oa* to integrate loan verbs from Spanish. This mechanism was borrowed into Yaqui, also Uto-Aztecan, where it is productive:

Ex. 14   NAHUATL (PASTORES) < SPANISH   (Wichmann & Wohlgemuth 2008)
ya   ki-puntaro-tika-ya   se   de   ihwante
now 3OBJ-point-PROG-IPFV one of them
'Now he had one of them at gunpoint.' [Sp. *apuntar*]

Ex. 15   YAQUI [UTO-AZTECAN] < SPANISH   (Wichmann & Wohlgemuth 2008)
che'a chúkula into te retratár-oa-wa-k
more later CONJ 1PL portray-oa-PASS-PERF
'And later we were photographed.' [Sp. *retratar*]

Languages with more than one strategy

Ex. 16   COPTIC (LYCOPOLITAN DIALECT)   < GREEK   (Grossman 2010a)
a-f-er-keleui   *a-f-keleui*
PAST.AFF-3SG.M-do-order   PAST.AFF-3SG.M-order
'He ordered'   'He ordered'
LIGHT VERB strategy   DIRECT INSERTION strategy

Ex. 17   SPANISH < ENGLISH   (Wohlgemuth 2009)
hacer clic   *clic-ar*
to do/make click   click-INF
'to click'   'to click'
LIGHT VERB STRATEGY   DIRECT INSERTION STRATEGY

Ex. 18   GREEK < ENGLISH   (Wohlgemuth 2009)
kán-o klik   *klik-ár-o*   (LVM = loan verb marker)
do-1SG click   click-LVM-1SG

'I click'   'I click'

LIGHT VERB STRATEGY   INIRECT INSERTION STRATEGY

Ex. 19   TURKISH < ENGLISH   (Wohlgemuth 2009)
              klik et-mek   *klik-le-mek*
              click do-INF   click-VERBALIZER-INF
              'to click'   'to click'
              LIGHT VERB STRATEGY   INIRECT INSERTION STRATEGY

The choice of a given strategy cannot be predicted based solely on the structural properties of the languages involved.

Hierarchy of verb borrowing strategies
(Wichmann & Wohlgemuth 2008, Wohlgemuth 2009)

LIGHT VERB < INDIRECT INSERTION < DIRECT INSERTION <| PARADIGM TRANSFER

'The degree to which a loan verb is integrated into the target language may be considered inversely proportional to the amount of formal mechanics expended by the target language on accommodating the loan verb (ignoring, for the present purposes, phonological aspects). From this point of view, the lowest degree of integration is associated with the light verb strategy, which involves a whole extra constituent for the integration. A somewhat higher degree of integration is associated with indirect insertion, where just an affix is required. In the case of direct insertion we have complete integration: here the loan verb is treated as if it were native. It is less straightforward to place paradigm transfer in the hierarchy, since on the one hand, no formal accommodation effort has been expended while, on the other hand, the loan verb is in a sense unintegrated inasmuch as it retains the inflectional morphology of the source language and resembles a code-switch. This accounts for our use of the composite symbol "<|", indicating that paradigm transfer might be considered part of the hierarchy, but that it has special status' (Wichmann & Wohlgemuth 2008).

… if a language has different patterns, these could correlate with the degrees to which the speakers of the target language are exposed to the source language(s) … more bilingualism would mean the choice of a strategy further to the right in the hierarchy and less bilingualism a strategy further to the left' (Wichmann & Wohlgemuth 2008: 109).

## 5. Evaluating predictions made by typological generalizations: the case of verb borrowing in the Coptic dialects

[…] All this makes it interesting to examine how various strategies are distributed among languages, what might motivate the choice of one strategy over another, and especially why the integration of verbs appears to be so complex in so many languages (Matras 2009: 177).

While quite a few languages have multiple strategies of verb borrowing, few have been studied in depth. Grossman (2010a) provides an examination of the Coptic dialects, all of which show considerable lexical borrowing from Greek.

איתן גרוסמן
טשטוטו78388מגע

Assumptions
- Language change is based on synchronic variation.
- A linguistic element **a** does not 'become' **b** in any simple sense. Rather, variants **a** and **b** co-exist in a given linguistic community; under certain circumstances, the relative frequency of variant **b** increases with respect to **a**.

Figure 1: language change as change in frequency of variants
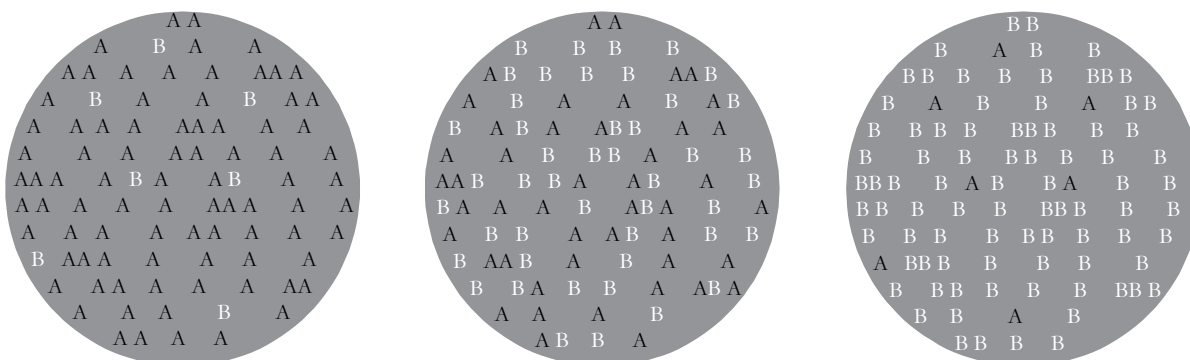
Predictions for the diachrony of verb borrowing strategies:
- Only adjacent strategies on the hierarchy will be attested as synchronic variants.
- In some cases, specific verb lexemes might be associated with a single strategy (lexical diffusion).
- Dialects that show an *overall* greater pattern of borrowing will prefer strategies to the right, i.e., with greater integration of borrowed verbs.
- Change will be towards the right (probably excluding the Paradigm Transfer Strategy).
- Change will probably involve (i) loss of light verbs (LIGHT VERB > DIRECT INSERTION) or (ii) the grammaticalization of LVMs (Loan Verb Markers) out of light verbs (LIGHT VERB > INDIRECT INSERTION).
- Another likelihood is the loss of specific verbal morphology from the source language, which cuts across strategies.

Greek and Coptic
- Uncertain sociolinguistic situation with respect to bilingualism.
- Barely and pre-Coptic lexical borrowing from Greek.
- Extensive lexical borrowing, up to and including adpositions.
- First borrowed verbs attested in very late Demotic, the stage of Egyptian immediately preceding Coptic. These verbs are borrowed as infinitives, governed by a light verb, and are written in Greek letters, in the opposite direction from the surrounding Demotic text.

a. The integration of Greek verb lexemes in Coptic: The traditional view

	Bohairic	**Sahidic**
Basic verbal construction	*a-f-sôtm*	*a-f-sôtm*
	PAST-3SG.M-hear[Coptic]	PAST-3SG.M-hear[Coptic]
	'He heard...'	'He heard...'
With denominal verb	*a-f-**er**-nobi*	*a-f-**r**-nobi*
	PAST-3SG.M-do-sin[Coptic]	PAST-3SG.M-do-sin[Coptic]
	'He sinned...'	'He sinned...'
With Greek verb lexeme	*a-f-**er**-keleuin*	*a-f-**ø**-keleue*
	PAST-3SG.M-do-order[Greek]	PAST-3SG.M-ø-order[Greek]
	'He ordered...'	'He ordered...'

b. Relevant Parameters:
- Presence or absence of light verb (*er-/ø*)   ±LM
- Presence or absence of Greek infinitive marker (*-n*)   ±IM

c. Patterns of Greek verb lexeme integration as actually occurring

	-n	*ø*
er-	(1a) *er-keleu̯in*	(2a) *er-keleu̯i-ø*
ø	(1b) *ø-keleu̯in*	(2b) *ø-keleu̯in-ø*

1a. *af-er-keleu̯i-n*   +LV, +IM (LV = Light Verb, IM = Infinitive Marker)
1b. *af-ø-keleu̯i-n*   −LV, +IM
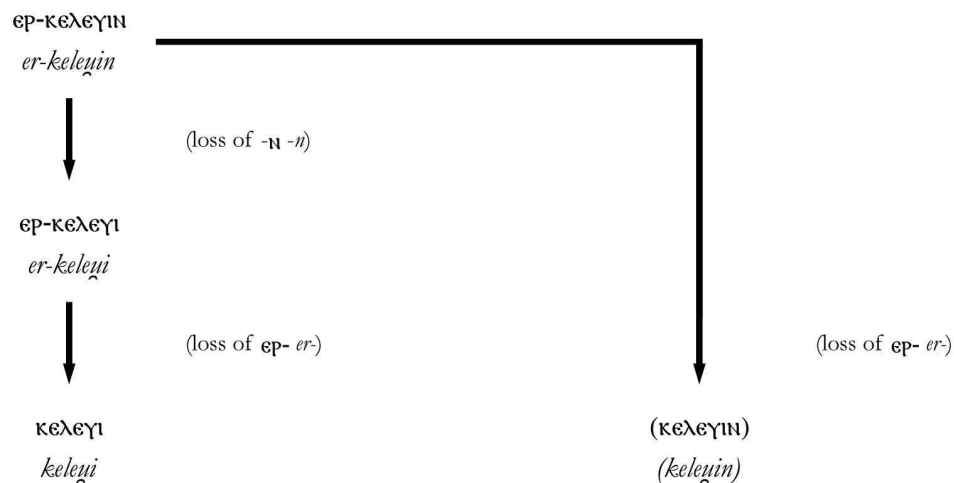2a. *af-er-keleu̯i-ø*   +LV, −IM
2b. *af-ø-keleu̯i-ø*   −LV, −IM

d. Pattern variation found in the Dialects
1. *er-keleu̯i-n* vs *er-keleu̯i*   +LV, +IM vs +LV, −IM
2. *er-keleu̯i-n* vs *er-keleu̯i* vs *keleu̯i*   +LV, +IM vs +LV, −IM vs −LV, −IM
3. *er-keleu̯i* vs *keleu̯i*   +LV, −IM vs −LV, −IM

e. Loan-verb integration patterns in the various Coptic dialects

	F7	N	B4	F4	F5	V4	S^nh	L4	L6	L*	L5	P	A	W	M	S
1a +LV, +IM	(+)	+	+	+	+	+	–	–	–	–	–	–	–	–	–	–
2a +LV, −IM	–	–	–	+	+	+	+	+	+	+	+	+	+	–	–	–
2b −LV, −IM	–	–	–	–	+	+	+	+		+	+	+	+	+	+	+

LESS INFLUENCE													MORE INFLUENCE			
CONSERVATIVE													INNOVATIVE			

f. Unidirectional pathways of change

ⲉⲣ-ⲕⲉⲗⲉⲩⲓⲛ
er-keleu̯in

(loss of -ⲛ *-n*)

ⲉⲣ-ⲕⲉⲗⲉⲩⲓ
er-keleu̯i

(loss of ⲉⲣ- *er-*)   (loss of ⲉⲣ- *er-*)

ⲕⲉⲗⲉⲩⲓ   (ⲕⲉⲗⲉⲩⲓⲛ)
keleu̯i   *(keleu̯in)*

g. The stages of integration of the loanword integration pattern

	Pattern 1a	Pattern 1b	Pattern 2a	Dialects at this stage
Stage 1:	*af-er-keleuin*			(F7), B4, N
Stage 2:	*af-er-keleuin*	*af-er-keleui-ø*		F4
Stage 3:	*af-er-keleuin*	*af-er-keleui-ø*	*af-ø-keleui*	F5, V4
Stage 4:		*af-er-keleui-ø*	*af-ø-keleui*	L4, L5, L6, L*, A, P, S^nh
Stage 5:			*af-ø-keleui*	S, M, W

Results

- The light verb strategy + Greek infinitive pattern is earliest.
- Within the framework of the light verb strategy, variants of the verbal lexeme without the infinitive marker of the Greek loan verb arise, leaving what looks like a bare verbal stem.
- At the next stage, variants without the light verb arise, leaving what looks like 'direct insertion' of a bare verbal stem (which also looks like an imperative).
- It appears that in the final-stage dialects, a 'direct insertion' pattern is conventionalized and can productively integrate new borrowings. These are the dialects in which *overall* borrowing is most extensive.
- This shows how multiple patterns of borrowing emerge over time, and how 'direct insertion' can arise from a 'light verb' strategy.
- There is some evidence, not presented here, that in some corpora, specific lexemes are associated with a single borrowing pattern.
- It is not necessary to assume, contra many scholars, that the borrowed verb form was an imperative. Rather, the borrowed verb form was an infinitive, which is much more common cross-linguistically.
- The distinction between 'light verb' and 'indirect insertion' might not be a necessary one (Matras 2009, but Wohlgemuth 2009 argues for it strongly).
- Incidentally, since Coptic is one of two languages considered to have borrowed imperatives as the 'input form,' this borrowing pattern should be reconsidered.

Explanations

- A functional explanation can be given for the loss of the Greek infinitive marker *-n*, as it is basically redundant (= economy).
- A non-structural, non-functional explanation must be posited for the loss of the light verb. Over time, borrowed verbs came to be more and more integrated into the Coptic lexicon (= 'more native'), and less formal mechanics were needed in order to integrate them into the Coptic verbal morphology. This has to do with speaker perception of the various components of the linguistic repertoire.

Additional evidence can be found in the integration of borrowed verb lexemes into native valency patterns, an important and hitherto completely unresearched domain:

Ex. 20 English *telephone* > Hebrew *telefon* > Hebrew *t.l.f.n* 'to call by telephone'

Hebrew *letalfen le-/el-*		Hebrew *lehitkašer le-/el-*	
tilfan-ti	*le-yossi*	*hitkašar-ti*	*le-yossi*
telephone:PST-1SG	to-Joe	telephone:PST-1SG	to-Joe
'I called Joe.'		'I called Joe.'	

tilfan-ti	*ela-v*	*hitkašar-ti*	*ela-v*
telephone.PST-1SG	to-3SG.M	telephone.PST-1SG	to-3SG.M
'I called him.'		'I called him.'	

This valency pattern is used for verbs of oral address and movement, rather than the 'default' accusative marker *et/ot-*:

Ex. 21
nišak-ti	*et*	*rona*	*nišak-ti*	*ot-a*
kiss:PST-1SG	ACC	Rona	kiss:PST-1SG	ACC-3SG.F
'I kissed Rona.'			'I kissed her.'	

Ex. 22   GREEK *diôkein* + ACC 'to pursue' > COPTIC *diôke nsa-*

SAHIDIC | BOHAIRIC
nere-n-ioudai-diôke	*nsa-Iêsous*	*na-u-cʰoci*	*nsa-Iêsous*
IMPF-the:PL-Jews-pursue	after-Jesus	PST-3PL-pursue	after-Jesus
'The Jews pursued Jesus.'		'They pursued Jesus.'	

Cf. Greek
ediôkon	*ton*	*Iêsoun*	*hoi*	*ioudaioi*
pursue:IMPF.3PL	the.ACC	Jesus.ACC	the.NOM	Jews.NOM
'The Jews pursued Jesus.'				

Rather than the 'default' accusative marker *n-/mmo-,* which is the valency pattern of most Coptic verbs of Greek or Egyptian origin:

Ex. 23
a-f-baptizei	*mmo-s*
PST-3SG.F-baptize	ACC-3SG.F
'He baptized her.'	

Rather than a group of segregated integration strategies of verbs, we find a continuum of devices. The theme of this continuum is not the structural effort that is made in order to accommodate verbs, but rather **the degree to which the 'verbness' of foreign-origin verbs is recognized and accepted**. Languges like Hebrew and Domari show that speakers apply considerable creativity and flexibility when it comes to the structural adaptation and integration of loan verbs […] [v]erbs accomplish, functionally speaking, two separate things. They are lexical signifiers that label events, activities, or states; and they also carry out the grammatical operation of anchoning the predication in the context of the utterance. On the far side of the verb integration continuum, we find languages that treat borrowed verbs merely as lexical labels, but do not entrust them with anchoring the predication. The latter task is instead delegated to a separate, 'light,' verb. […] At the other end of the continuum we find languages that recognize foreign verbs as both lexical labels (signifiers) and predicate-initiating devices. In between we find a variety of strategies (Matras 2009: 181-182).

In the Coptic dialects, we find a continuum of integration patterns ('devices') that show Greek loan verbs as lexical labels only (infinitives), and on the other, as both lexical labels and 'predicate-initiating devices.'

Why don't more languages use 'Paradigm Insertion'?

A final question: given the overwhelming prevalence of verb borrowing cross-linguistically, why don't speakers 'simply replicate loan verbs together with their original inflection'? (Matras 2009: 182). Matras gives several answers, which fit well with the fact that this phenomenon is almost entirely attested in populations with extensive and universal bilingualism (Romani, Domari, Neo-Aramaic, Quechua):

- Replicating the full verb requires significant bilingual proficiency. Borrowing is fundamentally a matter of the spread of innovations from bilingual to monolingual speakers. It might be difficult for monolingual speakers to learn the borrowed inflectional systems.
- Borrowing verbs along with their inflection creates a system with 'compartmentalized' inflectional subsystems, which could be very different from each other and therefore uneconomical or inefficient for speakers.
- It could be that the devices that encode predication are more symbolically representative of the distinction between the two parts of a multilingual repertoire.

Directions for future research
- Sociological-historical typologies of contact situations.
- Studies of integration strategies normally neglected, e.g., integration into valency patterns.
- Typological studies of borrowing of other categories.
- Typological studies of semantic borrowing ('pattern replication').

And, close to home:
A closer study of borrowing patterns in Semitic languages.

איתן גרוסמן
לשונות במגע
13.4.2010

Bibliography

Butterworth, B., Comrie, B. & Dahl, Ö, 1984. *Explanations for language universals*. Mouton.

Comrie, B,., 2004. 'Form and function in explaining language universals,' in: Butterworth et al. (1984) 87-103.

Croft, W., 2003. *Typology and universals (2nd edition)*. Cambridge University Press.

Dryer, M., 2008. 'Order of Subject, Object and Verb,' in: Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.), The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, chapter 81. Available online at http://wals.info/feature/81. Accessed on 2010-05-13.

Funk, W.-P., 1988. 'Dialects wanting homes: a numerical approach to the early varieties of Coptic,' in Fisiak, J. (ed) *Historical dialectology: regional and social*. Mouton de Gruyter.

Greenberg, J., 1966. 'Some universals of grammar with particular reference to the order of meaningful elements,' in: Greenberg, J (ed.), *Universals of grammar,* 73-113. MIT Press.

Grossman, E., 2010a. 'Grammatical variation and language change: the case of Greek verb lexemes in the Coptic dialects,' forthcoming in J. Cromwell & E. Grossman (eds.), *Beyond free variation: scribal repertoires in Old Kingdom to Early Islamic Egypt.* Oxford University Press.

Grossman, E., 2010b. 'Towards a typology of adposition borrowing,' forthcoming in: E. Grossman & T.S. Richter (eds.), *Language contact in Late Antique and Early Islamic Egypt.* Mouton de Gruyter.

Haspelmath, M., 2002. 'Functional and formal explanation in linguistics' (nine handouts), available online at http://email.eva.mpg.de/~haspelmt/papers.html.

Haspelmath, M., 2010. 'Comparative concepts and descriptive categories in cross-linguistic studies,' forthcoming in *Language* 86.

Haspelmath, M. & Tadmor, U., 2009. *Loanwords in the world's languages: a comparative handbook.* De Gruyter.

Haugen, E., 1950. 'The analysis of linguistic borrowing,' in: *Language* 26: 210-231.

Hyman, L., 1984. 'Form and substance in language universals,' in Butterworth et al. (1984) 67-85.

Matras, Y. & Sakel, J., 2007. *Grammatical borrowing in cross-linguistic perspective.* Mouton de Gruyter.

Matras, Y., 2009. *Language contact.* Cambridge University Press.

Moravcsik, E., 1975. 'Verb borrowing,' in: *Wiener Linguistische Gazette* 8: 3-30.

Moravcsik, E., 1978. 'Universals of language contact,' in: J. Greenberg (ed.) *Universals of human language*. Stanford University Press. 94-122.

Moravcsik, E., 2003. *Borrowed verbs.* Unpublished manuscript, cited in Wichmann & Wohlgemuth 2008.

Muysken, P., 2000. *Bilingual speech. A typology of code-mixing.* Cambridge University Press.

Ramat, P., 1987. *Linguistic typology.* Mouton de Gruyter.

Shopen, T. (ed.), 2007. *Language typology and syntactic description (2nd edition).* Cambridge University Press.

Stolz, T., Bakker, D., & Palomo, S. (eds.), 2008. *Aspects of language contact. New theoretical, methodological and empirical findings with special focus on Romanisation processes.* Mouton de Gruyter.

Thomason, S. & Kaufman, T., 1988. *Language contact, creolization, and genetic linguistics.* University of California Press.

Wichmann, S., 2004. 'Structural patterns of verb borrowing,' paper presented at the Wokshop on Loan Word Typology, Max Planck Institute for Evolutionary Anthropology, Leipzig 1-2 May 2004.

Wichmann, S. & Wohlgemuth, J., 2008. 'Loan verbs in a typological perspective,' in: Stolz et al. (2008): 89-121.

Wohlgemuth, J., 2009. *A typology of verbal borrowings.* Mouton de Gruyter.

Appendix: The Coptic dialects (Funk 1988, Grossman 2010a)

Akhmimic: all extant literary Akhmimic sources were used (see Funk 1985).

Bohairic: here refers to Early Bohairic (B4), a corpus of pre-conquest literary texts.

Fayyumic: a dialect continuum with much internal variation, each variety assigned its own siglum.

- **F4** An early variety of Fayyumic, attested in a small number of mss. The text used here is Crum and Kenyon 1900).
- **F5** 'Classical Fayyumic.'
- **F7** 'Proto-Fayyumic,' an especially archaic form of Fayyumic, with almost no Greek verbs or particles, and with but sparse Greek nouns.
- So-called **F56** is not a distinct dialect but is rather late (9th century) literary Fayyumic from Hamouli, a scriptorium in which dialect convergence between Fayyumic and Sahidic (see below) is visible in both the destandardized and localized 'Fayyumic' and 'Sahidic' texts.

Lycopolitan: four distinct dialects (Funk 1985)

- **L4** 'Manichaean' Lycopolitan
- **L5** Mainly Biblical mss.
- **L6** Mainly Gnostic mss.
- **L*** The dialect of Kellis, all texts found *in situ* in Ismant el-Kharab (Funk 1999, Shisha-Halevy 2002). Considered by Funk to be a *koinê*.

Middle Egyptian ('Oxyrhynchitic,' Mesokemic, see Funk 1981): a variety safely assignable to the Oxyrhynchos area. Here the texts consulted are Codex Scheide (Matt.) and (Acts).

Narmuthis: a language variety written in Demotic script, which is in fact much closer linguistically to Coptic varieties than to even the latest Demotic. Found in Medinet Madi. Only a small fraction of the 1500 ostraca has been published so far. All published ostraca have been consulted.

Palaeo-Theban: an especially archaic variety of Coptic, with close affinities to Akhmimic. Found in a single manuscript of Proverbs (P. Bodmer VI) .

Sahidic: the standard literary language of Egypt until its replacement by Bohairic sometime after the Muslim conquest of Egypt. The majority of Coptic texts are written in some form of Sahidic, often with significant regionalization. On Sahidic, its problematic linguistic and sociohistorical status, see Funk 1988: 152-154.

Other dialects without geographically-based names include:

- **I** 'Proto-Lycopolitan.'
- **I7** Closely related to **L** and **A** dialects, attested in three fragmentary texts.
- **K** A variety closely related to **B**, thought to come from Karanis. (P. Mich. 5421).
- **K7** is considered a variety of **K** (Kasser-Satzinger 1982, but see Funk 1988 for reservations).
- **V** With two subvarieties V4 (P. Mich 3520) and V5 (P. BM Or. 9271),
- **W** Closely related to Mesokemic. Found in a single manuscript (P. Mich 3521).
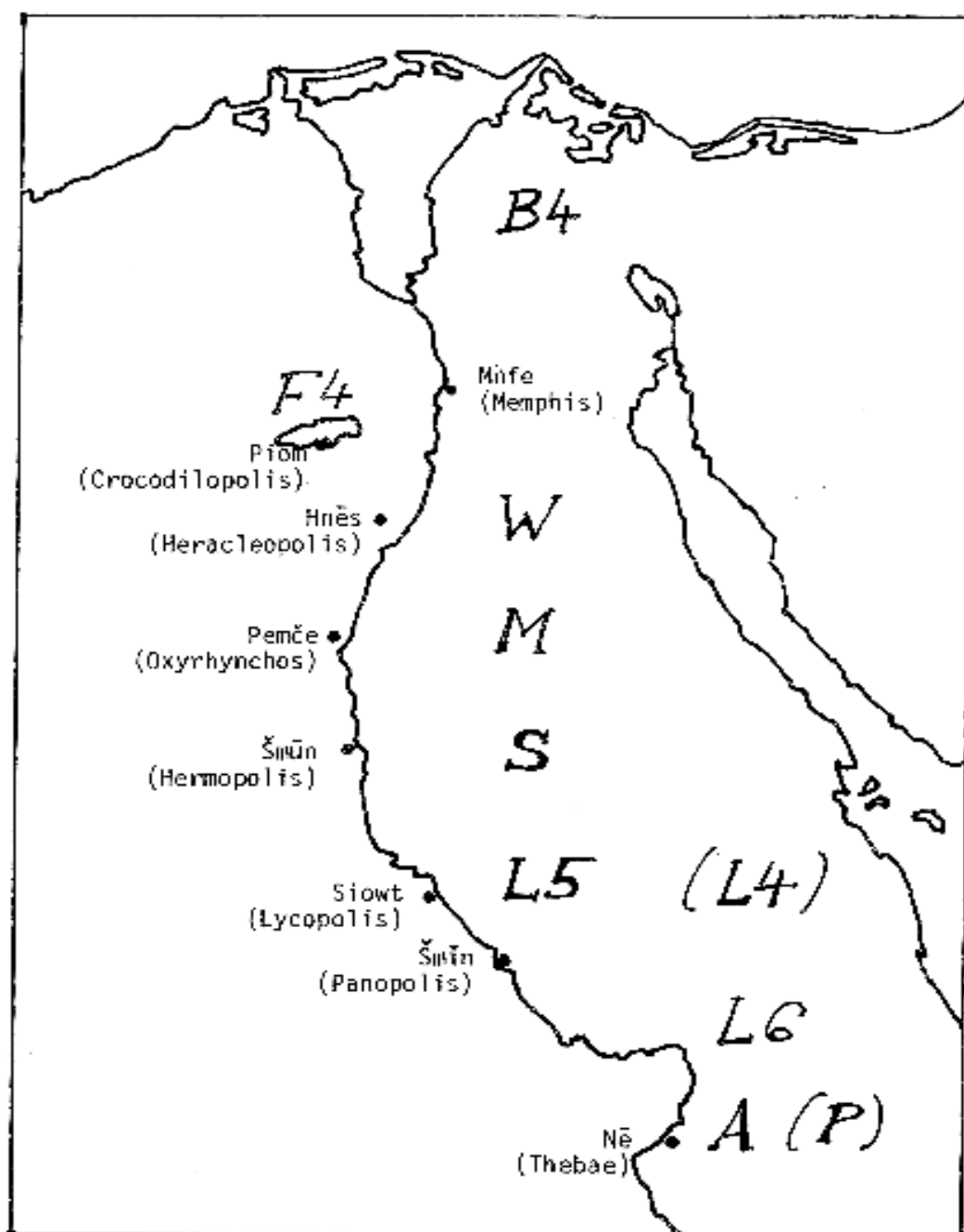
איתן גרוסמן
לשונות במגע
13.4.2010

Figure 6. Ten early Coptic literary dialects — a tentative plotting of their probable geographical affiliation according to the present results.