Cognitive contributions of translation practice to L2 performance

Robert M Maier
robert.m.maier@posteo.de
this talk

• translation in language teaching
• language processing (monolingual basics)
• bilingual language processing
• processing in translation tasks
• conclusion/outlook

• super-condensed summary:
  - translation practice helps to consolidate separate language representations
L2 teaching and translation

• grammar translation method
  – origins: e.g. Aelius Donatus (c4), *Ars grammatica*
    • teacher of rhetoric, grammarian
    • grammar as tool for teaching rhetorics
  – application to L2 teaching by mere extension?
  – criticisms of method (e.g. Brassai, 1881; Vietor, 1886)
    ~ criticism of lack of method


• grammars as phenomenological and/or classificatory writings to aid development of rhetorical skill

• language learner ~ student of rhetorics "with a handicap"

• texts in one language provide "something to say" in the other (i.e. L1>L2)
the success of GTM

- through implicit learning, through incidental learning, through...
- any particular advantages from translation?

- language processing
  - synaptic plasticity
    - cf presentations Grien, Langner

- "No matter what you do, they'll always pick up something in the long run"
  - reassuring inexperienced language teachers?
  - declaration of bankruptcy as a teacher?
  - declaration of bankruptcy as a researcher??
    - what next? "and then a mystery happens"? the legendary black box?

- why is translation still around in language teaching? pure stubborn traditionalism, or is there any real use to it?

- question of language processing and translation quite interesting, in terms of psycholinguistics: comprehension + production + non-native language

- synaptic plasticity: what is frequently activated gets easier to activate again; what is frequently activated together gets easier to activate together again
one of the first comprehensive models, 30 years old, still among the most feasible

modular, but strictly metaphorically so – think of this as a kind of primitive flowchart

synaptic plasticity not originally relevant, but more obviously so if connections between modules/nodes are understood to be subject to it
the meaning of models...

- subset hypothesis (Paradis 1987, 2004)
  - model ~ flow chart
  - no neurological correlate of "nodes"/"modules"

subset hypothesis: while connections exist between corresponding lemmas, lemmas are actually well entrenched (synaptic plasticity!) in activation contexts -- i.e. subsets of the lexicon that connect words of LA more strongly to other words of LA than to words of LB, and vice versa
in terms of **processing** in healthy individuals, image on left is sufficient

in terms of storage, image on right is more to the point:
- corresponding language features connected
- language-specific networks better connected in themselves than with each other
  - a consequence of synaptic plasticity, but (cf subsequent discussion Revised Hierarchical Model)

THIS image is too complicated to be helpful – consequence of combination of large-scale models (head) with small-scale features (words)
step 0: language in the mind

- a simplistic model

image on previous page implies intuitive assumption:

to talk about something, information is called from a memory node that connects concept with specific phonological speech plan

in comprehension, inversely, the incoming phonological signal is recognized as pertaining to a memory node that relates to a specific concept
however, comprehension progresses incrementally (not necessary for comprehension to hear entire word: "feath-" may be sufficient!)

[activation is always indicated with asterisks in this presentation]
step 1: incremental comprehension

- cohort model (Marslen-Wilson, ~1975; Taft & Hambly, 1986)
  - auditory input activates connected speech plans in onset cohorts
  - narrowing selection
- each speech plan (form) connected to a lexical entry (lemma)

speech plans group into co-activated cohorts by their identical onsets
selection narrowing down with every additional phoneme

distinction form/lemma
lemma level connects forms with concepts, and provides additional grammatical information (word class, grammatical gender, number, tense, suppletives, etc) – cf. Levelt, Roelofs & Meyer (1999)

img1: intake of first phoneme, selection of cohort with this onset
img2: intake of second phoneme, re-selection of cohort with this expanded onset
img3: intake of third phoneme (here, leading to point of uniqueness, recognition point)

real homophones may be disambiguated by context, e.g. "site"/"sight"
how does that work? hold on...
based on first incoming phoneme, initial cohort of likely comprehension candidates is selected (i.e. all candidates receive some activation)
based on second incoming phoneme, cohort of likely comprehension candidates is reduced (i.e. only matching candidates receive further activation)
etc., until recognition point reached (cohort reduced to only one candidate for recognition)

[fits in with Levelt's (1989) model, cf slide after references]
step 2: concept activation

- in comprehension

model presented here has multiple sources, might mention prototype semantics, might mention Levelt (1989) and Levelt & al. (1999)
activation of lemma/lexical node activates concepts that are represented by multiple semantic features
step 2: concept activation

- in comprehension
- in production
for production, semantic features are activated simultaneously, activation is passed on to concepts and lexical nodes they are connected to → node with maximum activation is selected to be used for production
step 2: concept activation

- in comprehension
- in production
- lexical priming

as shown in the cohort model, activation of nodes is not binary, but incremental; nor are nodes simply switched off after having been employed, but they slowly lose activation – which means there will be some activation left in what has been processed before, frequently helping to disambiguate e.g. input "site"/*sight"

...also leading to

lexical priming: residual activation from previous processing makes it also easier to access related concepts – e.g. "nurse" is produced more readily after "hospital", "doctor", "syringe" than after "revenue", "auditor", "tax"

(priming: processing of one cognitive concept is facilitated if stimulus processed last contains certain features, but not if it lacks these same features)
simplified illustration of (lexical) priming: after processing of word "nurse", residual activation in system makes comprehension of "hospital" easier

[similar phenomenon possible at level of speech plans (phonological priming), not illustrated here]
not all features are necessary to make a selection (of course)
even potentially relevant features may not be necessary (here e.g. in English)
step 3: structural processing

- morphosyntactic information stored with lemma (Levelt, Roelofs & Meyer, 1999)
- observation of syntactic priming (Bock, 1986)

reminder about grammatical (morphosyntactic) information (in production, plausibly contained in cognitive concept of state of affairs to be contained in utterance)

Bock (1986): observation of syntactic priming
- people tend to recycle the sentence type that they have just used
- passing freely between production and comprehension → challenge to earlier views of language processing
- not properly accounted for by Levelt's model
Pickering & Branigan 1998:
- report of lexical component in syntactic priming
  - syntactic priming is not increased by repetition of morphosyntactic categories (person, number, tense, aspect) (presumably mediated between lexicon and conceptualizer) but by lexical repetition ("lexical boost")
  - i.e. conceptualizer not involved in syntactic priming
  - plausible account: syntactic node connected to lemma!
from monolingualism to multilingualism

- so far: monolingual model
- bilinguals?

comprehension should fit into cohort model (impossibility to "switch" to another language in listening)
by and large works okay

insufficient representations may cause ambiguities
from monolingualism
to multilingualism

- so far: monolingual model
- bilinguals?
  - connecting system between language representations?
  - language choice: restricted by experience but not system; translation in children (Harris & Sherwood, 1978)
  - deBot (1992): separate formulators and conceptualizers, shared lexicon

no evidence for a connecting system between separate language representations e.g. through failures

evidence of connected systems: code-switching, translation translation NOT an acquired skill

de Bot: early bilingual transformation of Levelt (1989)
assumption of separate conceptualizers possible  
(situationally restricted experiences) but not necessary
from monolingualism to multilingualism

• so far: monolingual model
• bilinguals?
  – connecting system between language representations?
  – language choice: restricted by experience but not system; translation in children (Harris & Sherwood, 1978)
  – deBot (1992): separate formulators and conceptualizers, shared lexicon
  – Revised Hierarchical Model
developing bilingualism

- Revised Hierarchical Model (Kroll & Stewart, 1994)

as before: connections may (crucially) be subject to synaptic plasticity

process of SLA in RHM:
- stage 0: concepts + L1 knowledge (connected as in any native speaker)
- stage 1: new L2 items are associated with L1 items
- stage 2: developing connections L1>L2 permit active use of L2 items (used less frequently than L2>L1)
- stage 3: repeated use develops connections L2<>concept level (not used as frequently as L1<>concept level)

helps to account for **false friends**:
(L1 English, L2 German)

step 1: L2 speech plan "ernst" recognized as (passive) correspondence to L1 lemma serious
step 2: repeated experiences of using "ernst" as a correspondence to realize serious in an L2 context
step 3: development of L2-specific lemma to connect speech plan and concept level
SERIOUSNESS (NOT IN JEST)

serious

ernst

earnest

"serious"

"ernst"

"in earnest"

PHONOLOGICAL PROCESSING

outlined learning process is facilitated by pre-existence of (and thus easy access to) L1 speech plan "in earnest" (phonologically similar to new L2 speech plan, conceptually related)

(aka positive transfer)
new L2 speech plan shows surface (phonological) similarity with pre-existing L1 speech plan but conceptual dissimilarity →

"negative transfer", false friends
developing bilingualism

- Revised Hierarchical Model (Kroll & Stewart, 1994)
  - cf discussion: Brysbaert & Duyck (2010); Kroll et al. (2010)

RHM overly simplistic? (discussion ten years ago)

as shown, not necessarily to be understood modular; may reflect general relations between nodes pertaining to L1/L2
syntax and bilingualism

• Code-switching
  – Matrix Language Frame model (Myers-Scotton 1993):
    • lexical elements of one language can be embedded into morphosyntactic frame of another

• cross-linguistic structural priming (e.g. Hartsuiker, Pickering & Veltkamp, 2004):
  – syntactic information is capable of crossing language boundaries
  – formulators not "separate": shared structural nodes?

MLF plausible by Levelt’s model, too

CLSP another blow to that model, though (like monolingual structural priming)

lexical boost re-encountered as translation-equivalent boost (Schoonbaert et al., 2007)
...so what about translation proper?
(general model of cross-linguistic structural priming, here modified according to RHM for L1 German L2 English)

- residual activation of structural node accessible from all appropriate verbs
- translation-equivalent boost:
- access to residually activated structural node is moderated by connection with (also residually activated!) lexical node of other language
  - stronger support from residual activation of L1 lexical entry
  - weaker support from residual activation of L2 lexical entry

prediction: stronger priming in direction L1>L2 than L2>L1
translation

- **not** correlated with bilingualism
- not restricted to professionals – concomitant phenomenon with bi- and multilingualism (i.e. a "side effect")
  - Harris (1978): "we cannot teach how to translate, but only how to translate better"
- contested issue among professionals: directionality
  - quality of translation/interpreting L1→L2?
- own study

---

cf. multilingual abilities restricted by choice but not system:
- translation not possible without bilingualism
- translation capped by but not correlated to bilingualism (L2 proficiency no predictor of TL skill)
- translation develops concomitant with bilingualism (hence, even in very young children), but independent from it

axiomatic claim since early 1960s: better TL into L1

own study of priming in translation
study

• 32 participants
  – half L1 English L2 German, half L1 German L2 English
• task: translation German-English
• material: ditransitive constructions
  – Der Fürst zeigt dem Jungen das Flugzeug. (dative)
  – Der Fürst zeigt das Flugzeug dem Jungen. (inverted)
  – The prince shows the boy the plane. (DO)
  – The prince shows the plane to the boy. (PO)

participants: "regular" high-proficiency bilinguals without professional translator/interpreter training

task: spontaneous spoken translation into English of written German sentences ("sight translation")

materials:
• correspondence between all source and target sentences
• German verbs do not require prepositional construction (6/8 do not even permit)
• never (on any level) one single "correct" translation
PO is produced significantly more often by L1 German

PO is produced significantly more often after inverted stimuli

essential priming effect
- less pronounced for L1 German but nevertheless still significant
- priming strength between L1>L2 and L2>L1 hard to evaluate because of German PO preference
- also: measurement of response latencies
at first sight:
- responses as PO tend to be slightly faster
- responses to inverted by L1 German tend to be slower

but: two different types of input may play different roles in speakers' language representations (i.e. consider group differences per construction)
- datives: both groups get significantly faster over time responding with DO, but significantly slower responding with PO
- inverted: L1G get faster to produce PO, L1E get slower (tricky story because of preposition)

due to randomization, effect cannot hold sufficiently to become statistically significant straightforwardly, but present in the predicted way
priming in translation

- response choice: cued in both directions of translation, L1→L2 and L2→L1
- response time: reduced with reliable significance only for L1→L2
  - production of target (L2) response with construction parallel to source (L1)
  - advantage?
    - processing shortcut: residually activated parts of SL network employed in TL production
    - problem: employment (=activation) of SL network may open path to subsequent employment of SL features in TL production
- habituation may changed overall processing patterns (anecdotal evidence from T&I professionals)
- translation practice...
  - establishes representations of different languages as separate rather than shared processing entities
  - ...helps to avoid "Trojan syntax" (and perhaps more)

priming not absolute, but a strong tendency
outcome of processing is moderated by a multitude of factors

L2 responses generated
- faster with appropriate and source-like construction
- than with appropriate and target-only-like construction
- i.e. selection of cued construction employs residual activation in L1-specific connections (i.e. parts of L1 grammar)

problem: what is activated together often is activated together more easily, i.e. potential for continuing along grammatical paths of L1 (and thus ultimately of interpreting breakdown) is increased after cued responses!
Thank you.


Maier, R M (in preparation). The time course of primed and un-primed production in spontaneous translation.


models and their location in Levelt's (1989) model:

cohort model: between speech-comprehension system and lexicon

conceptual activation: between conceptualizer and lexicon (detour via formulator usually played down)

morphosyntactic processing with syntactic node: between conceptualizer, formulator and lexicon (regular morphosyntactic nodes related to grammatical encoding, syntactic node related to surface structure)
<table>
<thead>
<tr>
<th>L1 (16ea.)</th>
<th>source (16ea.)</th>
<th>...to DO</th>
<th>...to PO</th>
<th>...to other</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>DO</td>
<td>100</td>
<td>36</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>iDO</td>
<td>26</td>
<td>125</td>
<td>105</td>
</tr>
<tr>
<td>German</td>
<td>DO</td>
<td>72</td>
<td>117</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>iDO</td>
<td>37</td>
<td>162</td>
<td>57</td>
</tr>
</tbody>
</table>

- "other": incomplete reproduction, non-ditransitive structure, NP argument (subject, any object) substituted with "something/-body" etc.
- construction selection data from Maier, Pickering & Hartsuiker (2017)

- response latency data (Maier, in preparation):

<table>
<thead>
<tr>
<th>L1</th>
<th>source</th>
<th>...to DO</th>
<th>...to PO</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>DO</td>
<td>2793</td>
<td>2648</td>
</tr>
<tr>
<td></td>
<td>iDO</td>
<td>2658</td>
<td>2645</td>
</tr>
<tr>
<td>German</td>
<td>DO</td>
<td>2677</td>
<td>2709</td>
</tr>
<tr>
<td></td>
<td>iDO</td>
<td>3150</td>
<td>2835</td>
</tr>
</tbody>
</table>

my "translators" are really untrained multilinguals (i.e. regular language learners) (not even necessarily language students)