The idiom principle and the open choice principle

Holger Diessel
holger.diessel@uni-jena.de
The corpus linguist John Sinclair introduced the distinction between ‘the idiom principle’ and the ‘open choice principle’.

The open choice principle “is a way of seeing language text as the result of a very large number of complex choices. At each point where a unit is completed—a word or a phrase or a clause—a larger range of choice ends up, and the only restraint is grammaticalness.”
[from Moon 1998: 28]

The principle of idiom is that a language user has available to him a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments.
[Sinclair 1991: 110]
Definition of prefab:

A prefab is a combination of at least two words favored by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalization.

[Erman and Warren 2000: 31]

The most important criterion for the definition of a prefab is ‘restricted variability’:

By restricted exchangeability is meant that at least one member of the prefab cannot be replaced by a synonymous item without causing a change in meaning or function and/or idiomaticity.

[Erman and Warren 2000: 329]
Examples of ‘restricted variability’:

(1) \text{good friends} vs. \text{nice friends}
(2) \text{not bad [in the sense of ‘pretty good’]} vs. \text{not lousy}
(3) \text{I can’t see a thing} vs. \text{I can’t see an object}
(4) \text{I’m afraid I can’t help you} vs. \text{I am frightened I can’t help you}

Restricted syntactic exchangeability:

(1) \text{I guess} – (\text{* I don’t guess})
(2) \text{It will do} cannot lose its auxiliary (\text{*It does})
(3) \text{up here} cannot have reversed order (\text{* here up})
Prefabs can be divided into four basic types:

- Grammatical
- Functional
  - Grammatical
  - Pragmatic
  - Reducible
    - *in order to*
    - *waste time*
    - *You’re*
Prefabs are equally frequent in spoken and written discourse.

<table>
<thead>
<tr>
<th></th>
<th>Slots</th>
<th>Filled with prefabs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td>5,000</td>
<td>2,930 (58.6%)</td>
</tr>
<tr>
<td>Written</td>
<td>5,246</td>
<td>2,745 (52.3%)</td>
</tr>
<tr>
<td>Total</td>
<td>10,246</td>
<td>5,657 (55.4%)</td>
</tr>
</tbody>
</table>

But the proportions of the different types of prefabs vary across genres/modality.
Erman and Warren 2000

![Graph showing the comparison between spoken and written forms across lexical, grammatical, and pragmatic reducible categories. The graph indicates a higher prevalence of reducible forms in written compared to spoken speech.]
Erman and Warren 2000

![Bar chart showing lexical, grammatical, pragmatic, and reducible categories with spoken and written data.]

- **Lexical**: Spoken (30), Written (70)
- **Grammatical**: Spoken (20), Written (10)
- **Pragmatic**: Spoken (5), Written (5)
- **Reducible**: Spoken (0), Written (0)
Erman and Warren 2000

![Graph showing the comparison between spoken and written components of lexical, grammatical, and pragmatic reducibility.](image)
Erman and Warren 2000

![Bar chart showing lexical, grammatical, pragmatic, and reducible types in spoken and written contexts. The chart indicates a higher concentration of lexical reducibility in written compared to spoken contexts.]
### Average length of prefabs

<table>
<thead>
<tr>
<th>Type</th>
<th>Average Number of Words of Prefabs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td>2.61</td>
</tr>
<tr>
<td>Written</td>
<td>2.80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Average Number of Words of Prefabs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical</td>
<td>3.03</td>
</tr>
<tr>
<td>Grammatical</td>
<td>2.26</td>
</tr>
<tr>
<td>Pragmatic</td>
<td>2.29</td>
</tr>
<tr>
<td>Reducibles</td>
<td>2.00</td>
</tr>
</tbody>
</table>
# Lexical prefabs: Semantics

## Properties and states
- out of date
- be of help (to s.b.)
- different from s.th.
- have got s.th (=possess)

## Entities (abstract and concrete)
- sketch pad
- modern furniture
- permanent job
- intensive care

## Situations and events
- find one’s way (out of s.th.)
- go out of s.th
- it’s terrible weather
- make s.th. sure

## Places and positions
- here and there
- to the right
- in industry
- on paper

## Period or point of time
- at the time
- by then
- in the end
- a long/short time
Lexical prefabs: Formal

**Noun phrases**
rule of s.th.
great days of the past
a waste of time
the present state of our knowledge

**Verb phrases**
regard s.th. as s.th.
be in touch with
get the hand of s.th.
switch on the lights.

**Adjective phrases**
able to do s.th.
enough of s.th.
all right
suitable for s.th./s.b.

**Prepositional phrases**
for some reason
to the naked eye
on a clear night
in touch with s.th./s.b.

**Clause structures**
That doesn’t wash
Where have you been?
I’ve got to run.
Don’t worry (be happy).
# Emerging prefabs

## Emerging grammatical prefabs

<table>
<thead>
<tr>
<th>in reply to</th>
<th>provided that</th>
</tr>
</thead>
<tbody>
<tr>
<td>on the basis of</td>
<td>let alone</td>
</tr>
<tr>
<td>on these grounds</td>
<td>is going to</td>
</tr>
<tr>
<td>in response to</td>
<td>be just about to</td>
</tr>
<tr>
<td>in front of</td>
<td>due to</td>
</tr>
</tbody>
</table>

## Emerging pragmatic prefabs

<table>
<thead>
<tr>
<th>that’s true</th>
<th>I think</th>
</tr>
</thead>
<tbody>
<tr>
<td>it’s all my fault</td>
<td>I guess</td>
</tr>
<tr>
<td>that’s a good idea</td>
<td>I wonder (if)</td>
</tr>
<tr>
<td>what’s it all about</td>
<td>(Do you) remember</td>
</tr>
<tr>
<td>don’t bother</td>
<td>How do you know</td>
</tr>
</tbody>
</table>

Ich sag ma(l)
Lexical prefabs

Many lexical prefabs include slots:

- look forward to ___
- be in tune with ___
- bring ___ to a conclusion
- get in touch with ___
- take into ___ account

The slots of lexical prefabs are often semantically constrained:

- to waste time
- to waste effort
- to waste money
- to waste words
Lexical prefabs

Drive sb ___

- drive ___ crazy
- drive ___ insane
- drive ___ mad
- drive ___ up the wall

*drive ___ sane
*drive ___ happy
*drive ___ sad
*drive ___ depressed

We see this variability of prefabs as an outcome of two opposing forces constantly influencing language use: on the one hand there is a pull towards creating norms, facilitating communication, and at the same time there is a pull towards flexibility to allow expressibility, since expressibility increases with nonpredictability.

[Erman and Warren 2000: 41]
# Grammatical prefabs

Types of grammatical prefabs:

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determiners</td>
<td>of some kind, (the) kind of</td>
</tr>
<tr>
<td>Quantifiers</td>
<td>a bit (of), a little bit (of)</td>
</tr>
<tr>
<td>Proforms</td>
<td>each other, things like that, the kind of things</td>
</tr>
<tr>
<td>Introducers</td>
<td>there is, that is</td>
</tr>
<tr>
<td>Tense forming</td>
<td>be going to, have been</td>
</tr>
<tr>
<td>Aspect forming</td>
<td>begin to, be about to</td>
</tr>
<tr>
<td>Mood forming</td>
<td>could be, have got to, would rather</td>
</tr>
<tr>
<td>Prepositions</td>
<td>in front of, due to, on the basis of, on behalf of</td>
</tr>
<tr>
<td>Conjunctions</td>
<td>provided that, on condition that, let alone</td>
</tr>
<tr>
<td>Intensifiers</td>
<td>very much, more or less, if anything</td>
</tr>
</tbody>
</table>

Some grammatical prefabs have slots:

- On the basis of __
- There is __
- Of some kind of __
### Grammatical prefabs in German

<table>
<thead>
<tr>
<th>Category</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determiners</td>
<td>so ein</td>
</tr>
<tr>
<td>Proforms</td>
<td>der da</td>
</tr>
<tr>
<td>Introducers</td>
<td>es gibt</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>wird schon</td>
</tr>
<tr>
<td>Prepositions</td>
<td>in der Nähe von</td>
</tr>
<tr>
<td>Conjunctions</td>
<td>unter der Bedingung dass</td>
</tr>
<tr>
<td>Intensifiers</td>
<td>mehr oder weniger</td>
</tr>
</tbody>
</table>
Pragmatic prefabs

Text monitors
- Discourse markers: and finally, as I said, I mean, the thing is that
- Turn regulators: well you know, you see, well I think
- Repair markers: I mean, you know, well you know

Social monitors
- Interactives: hasn’t she? wouldn’t you? you see what I mean
- Feedback signals: I see, well no, suppose so
- Hesitations: what’s the word, I mean, you know
- Responses: yes I think so, well no, yes I see
- Performatives: thank you, good luck, why don’t you

Metalinguistic monitors
- Hedges: sort of, or somewhere, and all this
- Epistemological signals: I must say, I think, I should think
## Reducibles

<table>
<thead>
<tr>
<th>PRO am/is/are</th>
<th>I’m, You’re</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUX not</td>
<td>don’t, haven’t</td>
</tr>
<tr>
<td>AUX + AUX</td>
<td>would’ve, should’ve</td>
</tr>
<tr>
<td>let + us</td>
<td>let’s</td>
</tr>
</tbody>
</table>
Prefabs can be embedded into each other:

(1) [the preparations for [launching their rockets]]
(2) [the average of [forty miles [an hour]]]

Prefabs can also occur successively

(3) /I gather/ /(you’ve) been at it / /for nine years/

Prefabs and nonprefabs are often interspersed

(4) /that’s true/ -- --/it’s/ -- /a long time/
The discourse function of fixed expressions
[Moon 1998]

Holger Diessel
holger.diessel@uni-jena.de
Erman and Warren 2000

Lexical prefabs

Functional prefabs

Grammatical

Reducible

Pragmatic

Social monitors

Text monitors

Metalinguistic monitors
Moon 1998

- Informational
- Evaluative
- Situational
- Modalizing
- Organizational
Informational prefabs

(1) It was a great thrill to catch sight of my team mates as I got ....
(2) Bill is still in the running.
(3) The gunman opened the fire at close range from behind some foliage.

• Informational prefabs convey information, i.e. they describe processes, states, or qualities.
• They typically consist of a clause or verb phrase.
• 31 percent are metaphorical.

(1) behind bars
(2) by default
(3) face to face
(4) on sale
(5) clear one’s throat
Evaluative prefabs

(1) For him, this is only the **icing on the cake**.
(2) He is a **lame duck**.
(3) Bill remains **second to none**.

- Evaluative prefabs express the speaker’s opinions and evaluations.
- 47 percent are metaphorical.
- Many occur in copular constructions.

(1) do the trick
(2) over the top
(3) strike a balance
(4) down to earth
(5) get off to a flying start
Situational prefabs

(1) Hello, **good afternoon** Mr. Pit.
(2) **Excuse me**, but this is yours.
(3) **Thank you very much** for your help.

- Situational prefabs are typically found in spoken discourse.
- They usually refer to the extralinguistic context.
- They often function as independent speech acts.
- Syntactically, they are often expressed by full clauses.
- Some situational prefabs are old and deviate from grammatical rules.

(1) go for it
(2) good luck
(3) no problem
(4) so long
(5) long time no see
Modalizing prefabs

(1) These rules are by and large irrelevant.
(2) Believe you me.
(3) It stands to reason that …. 

• Modalizing prefabs indicate modality.
• Many modalizing prefabs have anomalous structures and are formulaic.

(1) as we know it
(2) in effect
(3) at any price
(4) I mean
(5) just trust me
Organizational prefabs

(1) **For examples**, ....
(2) **In spite of** the harsh conditions, I ....
(3) **Once upon a time** ....

- Organizational prefabs indicate links between linguistic elements.
- Some indicate links between elements of sentences (e.g. *in spite of, let alone*), others indicate links between larger chunks of discourse (e.g. *moreover, however*).
- Like modilizing prefabs, organizational prefabs are often syntactically anomalous and formulaic.

(1) all in all
(2) by the way
(3) let alone
(4) on the one hand
(5) with regard to
<table>
<thead>
<tr>
<th>Category</th>
<th>Function</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>informational</td>
<td>stating propositions, conveying information</td>
<td><em>rub shoulders with</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>in the running</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>catch sight of sth</em></td>
</tr>
<tr>
<td>Category</td>
<td>Function</td>
<td>Example</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>informational</td>
<td>stating propositions, conveying information</td>
<td><em>rub shoulders with in the running</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>catch sight of sth</em></td>
</tr>
<tr>
<td>evaluative</td>
<td>conveying speaker’s evaluations and attitudes</td>
<td><em>kid’s stuff</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>a different kettle of fish</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>it’s an ill wind</em></td>
</tr>
<tr>
<td>Category</td>
<td>Function</td>
<td>Example</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------------------------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>informational</td>
<td>stating propositions, conveying information</td>
<td><em>rub shoulders with in the running</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>catch sight of sth</em></td>
</tr>
<tr>
<td>evaluative</td>
<td>conveying speaker’s evaluations and attitudes</td>
<td><em>kid’s stuff</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>a different kettle of fish</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>it’s an ill wind</em></td>
</tr>
<tr>
<td>situational</td>
<td>relating to extralinguistic context, responding to situation</td>
<td><em>excuse me!</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>long time no see</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>knock it off!</em></td>
</tr>
<tr>
<td>Category</td>
<td>Function</td>
<td>Example</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------------------</td>
</tr>
</tbody>
</table>
| informational | stating propositions, conveying information  | *rub shoulders with in the running*
|             |                                               | *catch sight of sth*                        |
| evaluative  | conveying speaker’s evaluations and attitudes | *kid’s stuff*
|             |                                               | *a different kettle of fish*
|             |                                               | *it’s an ill wind*                         |
| situational | relating to extralinguistic context, responding to situation | *excuse me!*
|             |                                               | *long time no see*
|             |                                               | *knock it off!*                            |
| modalizing  | conveying truth values, advices, requests     | *you know what I mean at any price*
<p>|             |                                               | <em>needless to say</em>                          |</p>
<table>
<thead>
<tr>
<th>Category</th>
<th>Function</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>informational</td>
<td>stating propositions, conveying information</td>
<td><em>rub shoulders with in the running catch sight of sth</em></td>
</tr>
<tr>
<td>evaluative</td>
<td>conveying speaker’s evaluations and attitudes</td>
<td><em>kid’s stuff a different kettle of fish it’s an ill wind</em></td>
</tr>
<tr>
<td>situational</td>
<td>relating to extralinguistic context, responding to situation</td>
<td><em>excuse me! long time no see knock it off!</em></td>
</tr>
<tr>
<td>modalizing</td>
<td>conveying truth values, advices, requests</td>
<td><em>you know what I mean at any price needless to say</em></td>
</tr>
<tr>
<td>organizational</td>
<td>organizing text, signaling discourse structure</td>
<td><em>by the way talking of … be that as it may</em></td>
</tr>
<tr>
<td>Type</td>
<td>Frequency</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>informational</td>
<td>41%</td>
<td></td>
</tr>
<tr>
<td>evaluative</td>
<td>38%</td>
<td></td>
</tr>
<tr>
<td>situational</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>modalizing</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>organization</td>
<td>6%</td>
<td></td>
</tr>
</tbody>
</table>

Moon 1998