## Become Fluent in a Foreign Language by Using an Improved Technological Version of an Outdated Method

Michael Zock<sup>1</sup>, Guy Lapalme<sup>2</sup> & Lih-Juang Fang<sup>3</sup>

<sup>1</sup>Aix-Marseille Université (michael.zock@lif.univ-mrs.fr) <sup>2</sup>Université de Montréal (lapalme@iro.umontreal.ca) <sup>3</sup>Université Paris Sorbonne (fanglihjuang@yahoo.com)

Speaking a language can be an overwhelming task. The message (what to say), its corresponding linguistic expression (how to say it) and sound form (say it, i.e. articulation) have to be determined practically on the fly. To allow for this, parts of the process, in general the mechanical aspects (sentence structures) are automated, that is, they are carried out without requiring attention allowing the speaker to focus on other more important aspects, for example, the message (meaning). Speaking is a skill which has to be learned, and this can be supported in various ways. We propose an enhanced, computerized version of a well-known method: pattern drills. While being highly regarded in the sixties, pattern drills (PD) have become unpopular since then. Despite certain shortcomings we do believe in their virtues, at least with regard to the memorization of basic structures and the acquisition of fluency, the skill to produce language at a 'normal' rate. Of course, the method has to be improved, and we will show here how this can be achieved. Unlike tapes or books, computers are open media, allowing for dynamic changes, taking users' performances and preferences into account. Our drill-tutor, a small web application still in its prototype phase, allows for this. It is a free, electronic version of pattern drills, i.e. an exercise generator, open and adaptable to the users' ever changing needs

Journal of Cognitive Science 16-2: 151-173, 2015

Date submitted: 6/23/15 Date reviewed: 6/24/15-6/26/15

Date confirmed for publication: 6/28/15

<sup>©2015</sup> Institute for Cognitive Science, Seoul National University

#### 1. Problem

Language production is above all a goal-directed process where an author (speaker/writer) makes a set of choices (conceptual, linguistic) under specific space- and time constraints. Describing the nature of these choices and their consequences is the task of researchers: *linguists* (Fromkin, 1993), *psychologists* (Levelt, 1993, 1989) and *language engineers* (Reiter & Dale, 2000)<sup>1</sup>.

Speaking a language can be a daunting task. The message (*what to say*), its corresponding linguistic expression (*how to say it*) and acoustic form (*say it*) have to be determined on the fly. More precisely, we have to plan w*hat to say*, find the right *words* and the appropriate *sentence frame*. We have to *insert* then lemmas in the right place (syntax), *add* function words (syntax), perform the required morphological operations (agreement, ...) and while delivering the corresponding acoustic form (articulation), a complex task in its own right, plan the next conceptual fragment (conceptual level)<sup>2</sup>.

Next to speed, there is also a space problem. Messages can become complex, yet our short-term memory is limited (Miller, 1956; Baddeley, 1990). In other words, we cannot afford to plan everything in all its details before starting articulation. Our cognitive resources would simply be exhausted, causing us to forget what we meant to say, before having translated it into language. This is what has led researchers to build systems allowing for incremental processing: a lower component, say, the formulator, can start processing before the higher one, the conceptualizer, has fully completed its task (Kempen & Hoenkamp, 1987; de Smedt, 1990; 1994)<sup>3</sup>.

Obviously, language production is a complex process and so is its learning. What makes this process so difficult is the fact that the choices to be made are divers, numerous, and need to be carried out quickly. Speech is fast and its main components, words, have to be found on the fly in a huge repertoire, the mental lexicon (Zock, Ferret and Schwab, 2010)<sup>4</sup>.

Next, choices imply consequences, yet these are hard to predict. It

<sup>&</sup>lt;sup>1</sup>For a survey, see (Bateman and Zock, 2015).

 $<sup>^{2}</sup>$  For details concerning the various components, architectures and information-flow see de Smedt et al. (1996).

is nearly impossible to make a causal analysis based on input-output correlations. For once, a change of the meaning (conceptual input) may, though does not necessarily, lead to a change of the linguistic form. For example, gender differences of the grammatical subject are visible in English for the singular (He reads vs. She reads) but not for the plural. Also, the final output may depend on other choices than solely on the conceptual input (message). These may have been made in between these two extremes, for example, topicalisation, lexical choice, ....

Hence the connection between the input and output is not direct, it is generally mediated. The process is distributed over various levels (conceptual, lexical, morphological, ...), which implies that, even if the result of a choice is visible at the final output, its *causes* are not always easy to identify. While we can inspect the outcome of an action (or a choice) in a computer program, we cannot do the same with the human brain which is a black box. We can arguably see the input and the output, but not what happens in between. Actions (causes-consequences) occuring at the intermediate level are hidden. Put differently, in such cases we cannot tell what has caused what, neither can we make foolproof predictions concerning the consequences of a specific choice. For more details on this,

<sup>&</sup>lt;sup>3</sup> Before engaging in a conversation, a message must be planned. While there are many ways to perform this task, there are good reasons to believe that people do this in the following way: first an outline is planned (global, or skeleton planning), which is then filled in with details (local planning, elaboration). Planning proceeds thus from general to specific (depth first), that is, sentences are planned incrementally by gradual refinement of some abstract thought rather than in one go (one-shot process) where every element is planned down to its last details. Hence, eventhough we produce and perceive discourse as a string produced or analyzed left to right, it is planned from top-to bottom. This strategy is all the more likely as the message is going to be long. For details, see (Zock, 1996).

<sup>&</sup>lt;sup>4</sup> Consider the following. Normal discourse consists in 3-5 words per second. Suppose that of these 5 words 3 are content words. If a speaker were to consult a dictionary as an author might do, than this would imply something like 180 look-ups per minute in a lexicon, possibly containing 15000 words. Even in such a small-sized dictionary, this does not leave much time for finding the target. Word access can be compared to looking for the needle in a haystack, but under very severe time constraints.

see (Zock, 1994, 1988).

Let us briefly illustrate these points via two examples of French. The goal of the first (figure 1) is to show the multiple dependencies and consequences of a choice at the various levels (syntax, morphology, phonology). Imagine the following input [help (Paul, Marie)] which, at this stage is underspecified. In order to determine the final form we need additional information like, what is the topic, how to lexicalize a given concept, via what part of speech express a concept, etc. All these choices are likely to produce an effect, possibly visible though only at the internemdiate levels, and this is what we want to show. Hence, given our input and providing some additional information like, let us start the sentence with 'Paul (topic: Paul) Marie being known, we may get the following output: *Paul l'aide* (Paul helps her). Showing how a given choice (topicalisation, knowledge status of a given referent: old/new) affects the representation not only at the extremes (input-output), but also at the intermediate levels is a way to convert a black box into a glass box.

	Verb choice	Subject	pronoun	auxiliary	agreement
1	aider	Je (I)	les (them)	ai (have)	aidé <u>es</u>
2	venir en aide	Je (I)	leur (them)	suis (be)	venu en aide
3	porter secours	Je (I)	leur (them)	ai (have)	porté secours

Table 1: The far-reaching consequences of a simple choice

Above (table 1) is another example showing the far-reaching consequences of a choice. Suppose we were to express in French the following idea: [help <sub>past perfect</sub> (speaker, children)] where the two arguments (speaker, children) are known, hence likely to be expressed as pronouns. Let us see how a single choice, in our case the choice of the verb, may affect various aspects of the surface form. Since the notion of 'help' may be expressed in French via many verbal forms (aider, dépanner, venir en aide, porter du secours, donner un coup de main), we have to choose. Eventhough they are functionally speaking equivalent, they all are synonyms, each one of them dictates its specific constraints (type of object, preposition,...) or grammatical 'behavior'. Indeed, if you look at the table here below you'll notice changes concerning the case, affecting the form of the *pronoun* (les

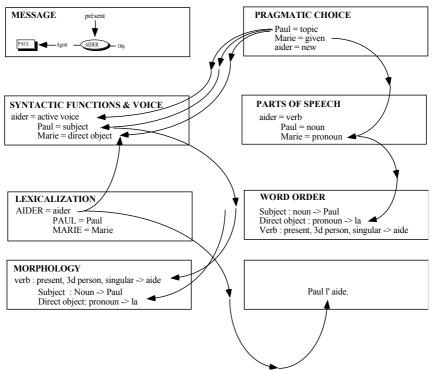


Figure 1: Language production represented as a glass box.

*vs.* leur), the nature of the *auxiliary* (être *vs.* avoir) and the object's *agreement* marked in the verb (aidée *vs.* aidé). In sum, a single choice may have various consequences, at various levels, and ultimately, be it only indirectly, for various elements of the surface form.

Next to *limited visibility* with respect to the consequences of a choice, there are *time-* and *space-management* problems. While there is certainly parallelism to some extent (Kempen & Hoenkamp, 1987), speaking is basically a sequential process, conceptualizing preceding formulation, formulation preceding articulation. Hence, any delay in one component, say, lexical choice, may yield a delay at the next lower component may need to be revised in the light of the results of a higher component allower component (retroaction). Correlated to the *time problem* (hesitations) there is also a

*space problem*. Any symbol or conceptual structure (concept/message) waiting for translation needs to be stored, taxing thus short-term memory, a very scarce resource (Miller, 1956).

As we have seen, producing language requires solving many tasks. Obviously, not all of them present the same level of difficulty, much depending also on the typological differences between the source and the target language. For example, *morphology* (cases, verb inflection,...) is known to pose a real problem for those learning German, Finish or Russian, while it is not really a challenge for languages like Chinese or Japanese. *Pronunciation* is considered to be an easy task by learners of Italian, Spanish or Japanese, while it seems to be a serious hurdle for those learning French or Chinese. This is not only due to the importance of tones<sup>5</sup>, but also to the number of unusual, yet similar sounding phonemes (zh, ch, sh, r, z, c, s, j, q, x) which can be a problem both for recognition (confusability) and pronounciability (production).

Hence, French may find it hard to distinguish and *hear* the differences between "zhi, zi, ji, chi, qi, xi, ci, si, shi, ri", since these sounds do not exist in their mother tongue. For similar reasons they will find it hard to produce the necessary sounds to convey a phone number like 47 17 74 [47  $\rightarrow$  (sì shí qī) 17  $\rightarrow$  (shí qī) 74  $\rightarrow$  (qī shí sì)]. Of course, French has its own set of challenges, for example, homophones and spelling. Students trying to learn French may find it hard to produce nasal sounds and their corresponding spelling: 'on', 'en' (*en*core), 'an' (*an*térieur), 'un' (*un* peu), 'in' (*in*grat).

In all these cases learners need to train their ears and tongues to understand or produce correctly a given sound or sound sequence in the target language. Yet they also have to learn to produce the corresponding graphical form. Chinese is known for the complexity of its writing system (ideograms : 马 or 妈 instead of the strings 'horse' or 'mother') which implies a steep learning curve, while French is reputed for its complex grapheme- sound mappings. For example, the following syllable /vɛ̃/ can be written in many ways<sup>6</sup>, all depends on the meaning.

In the remainder of this paper we will consider sentence production as a single task. Rather than focusing on it as a set of sub-tasks (syntax, morphology, lexical choice) we present it in a unified, manner. As we have seen, to speak freely and fluently, that is, to produce discourse spontanenously at a 'normal' rate is difficult. Yet, if speaking in one's mother tongue is already a daunting task, to do so in a foreign language is often overwhelming, at least at the initial stages (Bock, 1995). Language production is a skill (Levelt, 1975; de Keyser, 2007a) whose *elements* (words, rules, etc.) and *order* (what is to be processed and when) have to be learned, and this is hard work, requiring a lot of practice. This is nicely captured in the following precept or folk wisdom: use it or lose it.

#### 2. Goal, Scope and Related Work

Our goal is to help people to acquire quickly the skill of speaking. To this end we propose a relatively simple solution: the development of an open (i.e. customizable), generic, web-based environment. Put differently, we propose a digital version of "pattern drills" (Chastaing, 1969).

After having been very popular for many years, pattern drills and repetition which they rely upon have been discredited by linguists<sup>7</sup>, psychologists (Levelt, 1970, Herriot, 1971; Leont'ev, 1974; Krashen, 1981) and pedagogues (Rivers, 1964, 1972; Spolsky, 1966; Chastain, 1969; Savignon, 1983; Stevens, 1989, Wong and vanPatten, 2003).

While we partially agree with these criticisms with respect to the *creative aspects* of language production, we do not share them at all when *habit formation* or the *acquisition of automatisms* is the learning goal. Actually, it seems that we are not the only ones to hold this view, see for example, (deKeyser, 2007a, 2007b, 2001; Fitts, 1964; Garrod & Pickering, 2007; Gatbonton & Segalowitz, 2005, 1988; Guillaume, 1973; Hulstijn, 2001; Segalowitz, 2007, 2003, 2000; Segalowitz and Hulstijn, 2005), to name just these.

At least partial automatization of the process is necessary to become fluent in speaking. Automatisms are the speakers' means to buy time,

<sup>&</sup>lt;sup>5</sup>A given syllable, say 'ma', uttered with a different tone (mā; má; mǎ; mà), conveys a different meaning in Chinese: mā (妈) - mother ; má (麻) - hemp; mǎ (马) - horse; mà (骂) - to scold.

<sup>&</sup>lt;sup>6</sup> Vingt (twenty), vin (wine), vins (wines), il vint (he came), ...qu'il vînt (...that he came), je vaincs (I win), il vainc (he wins), en vain (in vain), ...

allowing them to focus on another, possibly more demanding component, for example, the next conceptual fragment, i.e. message, to be uttered. Put differently, in order to achieve the skill of fluent speaking, that is, fast conversion of ideas into sounds via language (Zock, 1997), we do believe that well-staged repetitions of stimulus-response patterns in a clear communicative setting, together with feedback are a valuable learning method. Of course, they are not the whole story. Interestingly enough, patterns have been rehabilitated by well-known linguists like Goldberg (1998) and by Ray Jackendoff (1993), one of Chomsky's most brilliant students.

Obviously, in order to be able to perform *automatically*, that is, without having to think about the various tasks mentioned, we must **exercise** them, as otherwise we will forget or be unable to integrate them into a well staged whole, a prerequisite for **fluency** (deKeyser, 2007).

In conclusion, we do believe in the virtues of PD, provided that the method is adapted and properly used. Also, languages need to be learned, and learning should be made simple and possible in a reasonably short time. Our goal is to help the learner reach the level of fluency needed to express his/her basic needs: ask for information, answer a question, solve a concrete problem, etc. by using the new language. In other words, our scope is the **survival level**. Yet, the more advanced learner probably needs other tools, though, even at that level patterns may have their role to play. Discourse also has structure (Hovy, 1993; McKeown, 1985; Zock, 1986). Just as words can be combined in specific ways to form a sentence, sentences can be combined to form a text. In both cases certain principles must be respected and these can be captured by patterns.

In sum, we propose the building of an exercise generator to help people to develop basic communication skills in a foreign language (for example, English, French, Japanese, Chinese). The goal is to assist the *memorization* of words and the acquisition of fundamental sentence patterns to become sufficiently *fluent* in the new language to participate in a simple conversation. Our approach is based on the following assumptions:

<sup>&</sup>lt;sup>7</sup> See for example Chomsky's violent criticism (Chomsky, 1959) of Skinner's book Verbal Behavior (Skinner, 1957)..

• **Resource limitations:** given the limitations of our brain (space and time), speakers cannot afford to perform very complex operations, especially not during learning;

• **Decomposition:** speaking being a complex process, we have to decompose it, allowing the speaker to focus selectively on a limited number of issues: *meaning, form,* or *sound*. Since people can focus only on few items at a time, it makes sense to put them into a situation, where they can rely upon a set of ready-made building blocks (the 'constants' of the pattern), computing only part(s) of the whole structure (the values of the patterns' variables).

• **Open-endedness:** different people have *different needs*. This being so, we propose to build an *open system*, allowing the user to tailor the tool to fit his or her needs.

• **Contextualization:** words are not learned in isolation, they are learned in the context of a sentence pattern, which may be part of even larger structures (paragraphs, discourse patterns).

• **Grounding:** words and sentence structures are linguistic resources used to achieve specific communicative goals. By indexing our patterns in terms of goals and by presenting words in the context of sentences, we achieve this kind of communicative grounding (pragmatic competency). The student learns when to use which specific resource.

• The need of practice: words have to be memorized, so do syntactic structures. Speaking is fast and various component tasks have to be carried out quasi-simultaneously. Hence we need to automate some of them (conversion meaning => form => sound). All these operations require practice (de Keyser, 2007), as without it we may not only forget, but also be unable to integrate the components into a well staged whole and to deliver the result in time.

• **Holism:** rather than assembling words into sentences we instantiate patterns. Instead of proceeding word by word, the learner operates on larger chunks, sentence patterns. In doing so, we gain what is needed next to knowledge, *space* (intentional resources) and *time*.

### 3. Some Theoretical Considerations Concerning our Approach

Since spontaneous language production is such a complex process one may wonder whether it can be taught or learnt. While languages can probably only learned and not taught, one can certainly try to make the learners' life easy. This is one of the main points in this paper.

Before presenting our approach, we would like to emphasize the following point. To go through the all steps depicted in Figure-1 and to apply rules assumed there or in natural language generation systems (Reiter & Dale, 2000) is highly unrealistic both for spontaneous discours (online processing) or language learning. There are various reasons justifying our doubts:

• *memory:* people do not have in their mind all the knowledge described by linguists, neither can they hold all the required information in their working memory (Baddeley, 1970);

- attention: people can focus only on a small set of items at a time;
- *time:* speech, i.e. the conception of a message and its translation into language is extremely fast. Speakers do not have the time to perform all the computations, i.e. search and apply the needed rules.

Linguists describe languages in terms of rules, but people hardly ever learn such descriptions, leave alone apply all of them, at least not at the initial stages of acquiring a new language. What people do learn though are *patterns* complying with these rules. Of course, people do use rules, but in conjunction with patterns.

Patterns can be seen as frozen instances of a given step in the derivational process described by generative grammars (Chomsky, 1957)<sup>8</sup>. They can be abstracted at any level of the process. They can be of any sort, even hybrid, mixing semantic and syntactic information. Patterns are global structures, which can be built dynamically by applying a set of rules. Moreover, they can be stored as ready made sentence plans or templates in which case they behave somehow like words: they can be retrieved at once, sparing us the trouble to have to go through the cumbersome process of structure creation. Obviously, access, i.e. pattern retrieval, is much faster than the

computation of its corresponding structure. There are simply too many steps involved. This is probably the reason why so many people use them for language (learners, interpreters, journalists, etc.) or for other tasks (music, programming, chess playing, etc.) without even being aware of it (Nagao, 1984).

Of course, there is a price to be paid: patterns need to be accessed (see below) and they may need to be accommodated. In other words, patterns have qualities, but also certain shortcomings: they are rigid and they tax long-term memory. Imagine someone abstracting a pattern for every morphological variation. Take for example the following two sentences: 'I've attended COLING 2010 in Bejing' *vs.* 'I plan to attend COLING 2016 in Osaka'. They basically instantiate the same pattern [(I've attended/ I'll attend <conference name> <time> <place>)]). In other words, it does not make sense in this case to abstract two patterns, since the two are so much alike. It would be much more reasonable to have one general pattern for the global structure and a set of parameters, i.e. rules for local adjustments, like agreement, tense, etc.

Just like patterns, rules have certain shortcomings. While they may account for the expressive power and all the regularities of a given language, they may prevent us from getting the job done in time, in particular if there are too many of them. This being so, we suggest to use a hybrid approach, resorting to each strategy when they are at their best, patterns for *global structures*, the syntactic layout, i.e. sentence frame, and rules for *local adjustments*. This combination gives us the best of both worlds, minimizing the use of computational resources (attention, memory), while maximizing the power (speed) and flexibility of output (possibly needed accommodations).

There is one problem though with this kind of approach. As the number of patterns grows, arises the problem of accessing them. This is where indexing plays a role. Patterns can be indexed from various points of view: *pragmatic* (communicative goals), *semantic* (domain), *syntactic*, etc. While we index our patterns pragmatically, i.e. in terms of communicative goals (function that the pattern is to fulfill), we allow their access also via other

<sup>&</sup>lt;sup>8</sup> https://en.wikipedia.org/wiki/Parse\_tree

means: navigation in a goal hierarchy.

Let us now briefly present our method.

### 4. Method

Learning requires some kind of support. While there are certainly many good teaching methods, there is at least one point where nearly all of them (books, tapes) fall short: due to the media constraints, they are closed. In consequence, everything has to be anticipated, which implies that all students have to take the same route, in spite of the diversity of their everchanging needs. This is a pitfall we try to avoid in our sentence and exercise generator, an open, customizable, web-based tool designed for novices studying foreign languages. The generator's **inputs** are *communicative goals* and *conceptual information*, the **output** is text (i.e. written form) or synthesized speech.

To support fluency acquisition, we suggest to build a *template-based* sentence generator. Patterns or templates are abstractions over concrete linguistic instances, i.e. sentences (I prefer beer to wine => < SPEAKER > prefer < DRINK<sub>1</sub> > to < DRINK<sub>2</sub> >). Patterns are linked to *communicative* goals, for example, 'comparison', the speaker's starting point (see table 2 and figure 2).

	goal	associated pattern	example of instantiated pattern
1°	Identity (name)	My name is <name>, <first name=""> <last name="">.</last></first></name>	My name is Bond, James Bond.
2°	Presentation (full name)	This is <first name=""> <last name=""> also called the <sur-name>.</sur-name></last></first>	This is <i>Bjorn Borg</i> , also called the <i>iceberg</i> .
3°	Origin (country)	I am from <country> and you?</country>	I am from Portugal, and you?
4°	Q-A: preference	Q: What do you like better <drink1><drink2>.</drink2></drink1>	Q: What do you like better, <i>tea</i> or <i>coffee</i> ?
		A: I prefer <drink1> to <drink2>.</drink2></drink1>	A: I prefer tea to coffee.
5°	Q-A: comparison	Q: Which city is bigger, <place1> or <place2>?</place2></place1>	Q: Which city is bigger, Tainan or Taipei?
		A: <place1> is bigger than <place2>.</place2></place1>	A: Taipei is bigger than Tainan.

Table 2: Patterns indexed in terms of goals

Note that we start from the assumption that the user knows the goal s/he would like to achieve. Hence, the dialogue goes as follows (see figure-2). Given some goal (step-1), the system presents a list of patterns from which

the user must choose (step-2), and instantiate then the pattern's variables with lexical items (step-3) and morphological values (step-4, or the steps A-D in figure 2). Note that some of these choices could be considered as optional, as they are performed by the system. This is typically the case in traditional pattern drills, where the user has no choice whatsoever concerning the input, the words to use.

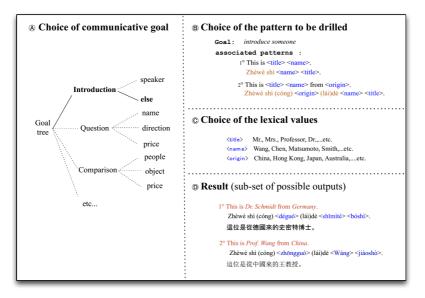


Figure 2: Communication Flow

Note also that the *conceptual input* (see table-3) is distributed over three layers: at a global level (macro-level) the speaker chooses the pattern via a *goal* by providing incrementally *lexical values* (for the pattern's variables) and *morphological parameters* (number, tense) to refine gradually the initially underspecified message. This kind of distribution has several advantages. Information is requested only when it is relevant and needed. There is a better control in terms of access, storage and processing load. Obviously, this approach is better than storing a pattern for every morphological variant. Last, but not least, this method is faster for conveying a message than navigating through a huge *lexical* or *conceptual* 

*ontology,* as suggested elsewhere (Zock,1991; Power et al., 1998; or Zock et al. 2009).

task	input	output
MACRO-LEVEL 1) choice of goal		set of sentence frames
		$\begin{array}{l} 1 <\!\!\mathrm{OBJECT_1}\!\!> is more <\!\!\mathrm{ATTRIBUT}\!\!> than <\!\!\mathrm{OBJECT_2}\!\!> \\ 2 <\!\!\mathrm{OBJECT_2}\!\!> is less <\!\!\mathrm{ATTRIBUT}\!\!> than <\!\!\mathrm{OBJECT_1}\!\!> \end{array}$
	2) choice of <i>sentence frame</i>	<object-1> is more <attribut> than <object<sub>2&gt;</object<sub></attribut></object-1>
MICRO-LEVEL	3) choice of <i>lexical value</i>	lexically specified structure
		Cigar is more expensive than perfume.
	4) morphological parameter	morphologically specified structure <object<sub>1&gt;: plural</object<sub>
		Fully specified conceptual, syntactic and morphological structure

Table 3: Conceptual input as a four-step process for the following output: *"Cigars are more expensive than perfume"* 

# 5. Building and Using the Resource

There are two aspects to be considered: building the resource and using it.

### 5.1 Building the Resource

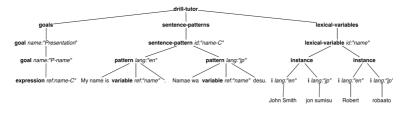
To **build the resource** (construction phase), we index a list of fundamental sentence *patterns* with *goals* from which the learner will choose during the exercice phase. Since different people have different needs we keep the system open so that the user can customize it according to his or her needs. In other words, the user can change certain parameters:

- the link between *patterns* and *goals*;
- the names of the goals (if s/he doesn't like our metalanguage);
- the words with which s/he'd like to instantiate a given pattern;
- the *number of times* s/he'd like to work on a given pattern;
- the *time delay* between a stimulus (question) and a response (answer);

• and so on.

Given the model described above, we wanted to clearly separate the description of the goals and the means, i.e. sentence patterns, to achieve them. Furthermore sentence patterns will be generalized in terms of variables standing for different words chosen according to a particular situation. For example, to achieve the goal of presenting oneself, one can use one of the following sentence patterns "They call me X" or "My name is X" in which the lexical variable X can be instantiated by the real name of the speaker. The allowed values of these variables can be defined in a dictionary. For a given goal, corresponding patterns and lexical variables can be defined for many languages.

Given the hierarchical nature of the goals and the flexibility needed for representing them, we decided to use an XML structure to keep the information. With an appropriate editor, it allows a linguist to define new goals, patterns and dictionary entries without dealing with the intricacies of the computer program that interprets this structure to display it to the user. Figure 3 shows the tree and the corresponding XML structure for a simple pattern in two languages (English and Japanese).



<?xml version="1.0" encoding="UTF-8"?> <drill-tutor>

```
<goals>
<goal name="Presentation">
<goal name="P-name">
<expression ref="name-C"/>
</goal>
</goal>
```

```
<sentence-patterns>
```

```
<sentence-pattern id="name-C">
My name is <variable ref="name"/>.
Namae wa <variable ref="name"/> desu.
</sentence-pattern>
</sentence-patterns>
```

#### <lexical-variables>

```
<lexical-variable id="name">
<instance>
<i lang="en">John Smith</i><i lang="jp">jon sumisu</i>
</instance>
<i lang="en">Robert</i><i lang="jp">robaato</i>
</instance>
</lexical-variable>
</lexical-variable>
</drill-tutor>
```

Figure 3: Definitions of *goals, sentence patterns* and *dictionary entries* in tree form and in the corresponding XML.

#### 5.2 Using the Resource

The user chooses the goal of communication. The goals being treestructured, the user can either drill down to a given goal by clicking on a goal name to expand the sub-goals until a pattern is displayed with variables shown in bold. Figure 4 show the values of a lexical variable, here, nationality.

Goals being indexed it is also possible to search for them via a term. In Figure 5, the user asked for all goals and patterns dealing with the term "name". Having selected to exercise a given language, the user is shown the two steps displayed in Figure 6.

#### 6. Discussion and Conclusion

We have started the paper by stressing the fact that speaking in a foreign language is difficult. We have tried then to show that the acquisition of this skill could be made feasible by blending an old theory into new technology.

# **Drill Tutor Goals and Correspondances**

Welcome Li,

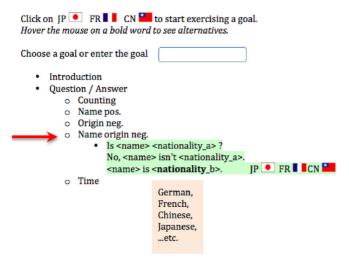


Figure 4: Lexical choice

Drill	Tutor Goals and Correspondences
Welcome (	- Guy
	FR, JP or CN link to start exercising a goal.
Hover the r	nouse on a bold word to see alternatives
Search for	a goal name
Search for	a goar name
• Prese	entation
0	name
	<ul> <li>My name is name. JP FR CN</li> </ul>
	<ul> <li>They call me name. FR CN</li> </ul>
0	origin
0	title name origin
	<ul> <li>This person is title name from country. <u>JP FR CN</u></li> </ul>
	tion Answer
	Counting
0	name pos
	<ul> <li>Are you name_a? Yes, I am name_a. <u>FR JP CN</u> origin neg</li> </ul>
	name origin neg
	<ul> <li>Is name nationality_a ? No, name isn't nationality_a. name is nationality_b. FR JP CN</li> </ul>
0	time

Figure 5: Search and display all goals involving the name of a person. After

each pattern are displayed the languages (links) in which one can carry out these exercice. JP: Japanese; FR: French; CN: Chinese

00	Presentation - name		
Your goals : Presentation - name			
Guy, translate th	Guy, translate this sentence into Japanese		
My name is Robe	My name is <b>Robert</b> .		
Hit return and compare	with our proposed answer		
00	Presentation - name		
Your goals : Presentation - name			
Guy, translate this sentence into Japanese			
My name is <b>Robert</b> .			
Namae wa roba	ato desu.		
なまえ わ ろばあと です.			
Did you get it right? yes no			
Select a different goal			

Figure 6: The top shows the initial screen with an instantiated sentence pattern. The user is invited to translate the prompted sentence in the chosen language and, after typing a carriage return; she is shown the proposed answer and asked if her answer was right. The number of correct or wrong answers to each goal is saved by the system and the user is shown another instance of the same pattern. It is also possible to choose another pattern (see Figure 5).

While the current prototype is fairly small (15 goals, a dozen of patterns and 300 words, in four languages), this should not be taken as an argument against the potential usefulness of our approach. Our focus was not on coverage but on generality. We wanted to see how difficult it would be to use the very same approach for typologically different languages. We were pleasantly surprised to see that even adding Chinese, after having built the system for French and Japanese, was quite simple and this certainly holds for most if not all languages devoid of a complex morphology.

In sum, the number of patterns and the size of the vocabulary is not really our major concern at this stage, the focus being on the implementation of an editor designed for building, modifying and using a database. The database can easily be extended. Note also that our system is not only an exercise generator, but also a language generator, simple as it may be. Our drill tutor has several features that set it apart from traditional pattern drills, user-controlled input being just one of them.

We have argued that producing sentences by applying only rules was too cumbersome, and that patterns alone were too rigid and too greedy concerning storage (memorization) and access, hence we've settled for a compromise, using patterns for the global structure and rules for the local refinements. This improves speed considerably, while minimizing storage of patterns and rules.

To conclude, we do believe in the virtues of our approach: the system is open and customizable with respect to input, linguistic knowledge, processing preferences, interface, etc. It is generic and it can be built and extended quite easily, by allowing to add various plug-ins: synthesized speech (certainly an asset or even a must for languages like Chinese or French), automatic creation of patterns or automatic building of a pattern library. Obviously, the ultimate judge of the qualities of the system is the user, but since we are still in the development phase, this is left for the future.

It is clear that pattern drills are not a panacea. They can even be harmful if not used properly (parroting, mindless repetition), but used in the right way, which means, at the right moment, with the right goals and at the right proportion, they can do wonders. Just like a tennis player might want to go back to the court and train his basic strokes, a language learner may feel the need to drill resisting patterns. Whoever has tried to become skillful in a language fundamentally different from his own can't but agree with deKeyser's (2001) words when he writes: "Without automatization no amount of knowledge will ever translate into the levels of skill required for real life use".

#### References

- Baddeley, A. (1990). *Human memory: theory and practice*. Hillsdale: Erlbaum.
- Bateman J & M. Zock. (2015) Natural Language Generation. In R. Mitkov (Ed.) Handbook of Computational Linguistics (2nd edition), Oxford University Press. Forthcoming.
- Bock, J.K. (1995). Sentence production: From mind to mouth. In J. L. Miller& P.D. Eimas (Ed.), *Handbook of perception and cognition*. Vol. 11:Speech, language and communication. Orlando, FL: Academic Press.
- Chastain, K. (1969). *The audio-lingual habit learning theory vs. the code-cognitif learning theory*. IRAL, 7(2), 97–107.
- Chomsky, N. (1959). A review of B. F. Skinner's Verbal Behavior. Language, 31(1), 26-58.
- Chomsky, N. (1957). Syntactic Structures. The Hague, Mouton
- deKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). New York: Cambridge University Press.
- deKeyser, R. (2007a). Skill acquisition theory. In J. Williams & B. VanPatten (Eds.), *Theories in Second Language Acquisition: An introduction* (pp. 97-113). Mahwah, NJ: Erlbaum.
- deKeyser, R. (Ed.). (2007b). *Practicing in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press
- de Smedt, K. (1990). IPF: An incremental parallel formulator. In R. Dale, C. Mellish & M. Zock (Eds.), *Current research in natural language* generation (pp. 167–192). London: Academic Press.
- de Smedt, K. (1994). Parallelism in incremental sentence generation. In G. Adriaens & U. Hahn (Eds.), *Parallel natural language processing* (pp. 421–447). Norwood, NJ: Ablex.
- de Smedt, K., Horacek, H. & Zock, M. (1996). Architectures for Natural Language Generation : Problems and Perspectives. In G.Adorni & M. Zock (Eds.), *Trends in Natural Language Generation: an Artificial Intelligence Perspective*, Springer Verlag, New York, pp. 17-46
- Fitts, P. (1964). Perceptual motor skill learning. In A. Melton (Ed.), *Categories of human learning* (pp. 243–285). New York: Academic press.
- Fromkin, V. (1993). Speech Production. In Berko-Gleason, J. & N. Bernstein Ratner, Eds. *Psycholinguistics*. Fort Worth, TX: Harcourt, Brace,

Jovanovich.

- Garrod, S. & Pickering, M.J. (2007). Automaticity in language production in monologue and dialogue. In A.S. Meyer, L.R. Wheeldon, & A. Krott (Eds.), *Automaticity and control in language processing* (pp. 1-21). Hove: Psychology Press.
- Gatbonton, E. & Segalowitz, N. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly*, 22, 473-492.
- Gatbonton, E. & Segalowitz, N. (2005). Rethinking communicative language teaching: a focus on access to fluency. *Canadian Modern Language Review*, 61, 325-353.

Guillaume, P. (1973). La formation des habitudes. PUF, Paris.

- Herriot, P. (1971). *Language and Teaching: A Psychological View*. London: Methuen and Co.
- Hovy, E. (1993). Automated Discourse Generation using discourse structure relations. *Artificial Intelligence*, 63, 341-385.
- Hulstijn, J.H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge, UK: Cambridge University Press.
- Jackendoff, R. (1993). *Patterns in the Mind: Language and Human Nature*. London: Harvester-Wheatsheaf.
- Kempen, G. & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201–258.
- Krashen, S. (1981). Second Language Acquisition and Second Language Learning. http://www.sdkrashen.com/ SL\_Acquisition\_and\_Learning/ index.html
- Leont'ev, A. (1974). *Psycholinguistik und Sprachunterricht*. Stuttgart, Germany: Kohlhammer.
- Levelt, W. (1989). Speaking. MIT Press, Cambridge, Mass.
- Levelt, W. (1993) The architecture of normal spoken language use. In Blanken G., J. Dittmann, H. Grimm, J. Marshall & C. Wallesch (eds.) *Linguistic Disorders and pathologies*. W. de Gruyter, Berlin, New York
- Levelt, W. (1975). Systems, skills and language learning. In A. van Essen & J.P. Menting (Eds.), *The Context of Foreign Language Learning* (pp. 83-99). Assen: Van Gorcum.
- McKeown, K. (1985). Discourse Strategies for Generating Natural-Language Text. *Artificial Intelligence*, 27, 1-41.
- Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological*

Review 63 (2): 81–97.

- Nagao, M. (1984). A framework of a mechanical translation between Japanese and English by analogy principle. In A. Elithorn & R. Banerji (Eds.), *Artificial and Human Intelligence* (pp. 173–180). Amsterdam: Elsevier.
- Power, R., Scott, D. & Evans, R. (1998). What you see is what you meant: direct knowledge editings with natural language feedback. In H. Prade (Ed.), 13th European conference on artificial intelligence (ECAI'98) (pp. 677–681). Chichester: Wiley.
- Reiter, E. & Dale, R. (2000). *Building natural language generation systems*. Cambridge: Cambridge University Press.
- Rivers, W. (1964). *The Psychologist and the Foreign Language Teacher*. University of Chicago Press,.
- Rivers, W. (1972). Speaking in Many Tongues: Essays in Foreign Language Teaching. Rowley: Newbury House.
- Savignon, S. (1983). Communicative Competence: Theory and Classroom Practice. Reading, MA: Addison-Wesley.
- Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 200-219). Ann Arbor, MI: University of Michigan Press.
- Segalowitz, N. (2003). Automaticity and second language learning. In C. Doughty & M. Long (Eds.), *The handbook of second language* acquisition (pp. 382-408). Oxford, UK: Blackwell.
- Segalowitz, N. (2007). Access fluidity, attention control, and the acquisition of fluency in a second language. *TESOL Quarterly*, 41, 181-186.
- Segalowitz, N. & Hulstijn, J. (2005). Automaticity in second language learning. In J. Kroll & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). New York: Oxford University Press.
- Skinner, B. F. (1957). Verbal behavior. New York: Appleton-Century-Crofts.
- Spolsky, B. (1966). A psycholinguistic critique of programmed foreign language instruction. *IRAL*, 4(2).
- Stevens, V. (1989). A direction for CALL: From behavioristic to humanistic courseware. In M. Pennington (Ed.), *Teaching Languages with Computers* (pp. 31–43). La Jolla, CA: Athelstan.
- Weir, R. (1962). Language in the Crib. Mouton, La Haye.
- Wong, W. & VanPatten, B. (2003). The Evidence is IN: Drills are Out. *Foreign Language Annals*, 36(3), 403-23.
- Zock, M. (1997). Sentence Generation by Pattern Matching: the Problem of Syntactic Choice. In R. Mitkov & N. Nicolov (Eds.), *Recent Advances*

in Natural Language Processing. (pp. 317-352).

- Zock, M. (1996). The Power of Words in Message Planning. 16th International Conference on Computational Linguistics, Copenhagen, pp. 990—995.
- Zock, M. (1994) Language in action, or, learning a language by watching it work. 7th Twente Workshop on Language Technology: Computer-Assisted Language Learning, Twente, pp. 101-111
- Zock, M. (1991). SWIM or sink: the problem of communicating thought.
   In M. Swartz & M. Yazdani (Eds.), *Intelligent tutoring systems for foreign language learning* (pp. 235–247). New York: Springer.
- Zock, M. (1988) Natural languages are flexible tools, that's what makes them so hard to explain, to learn and to use. In Zock, M. & Sabah, G. (eds.), *Advances in Natural Language Generation: an Interdisciplinary Perspective*, Ablex, Norwood, N.J., Vol. 1, pp. 181-196.
- Zock, M. (1986) The read thread or text grammars as guides for organising and expressing thoughts in the mother tongue or a foreign language (original in French: Le fil d'Ariane : ou les Grammaires de Texte comme Guide dans l'Organisation et l'Expression de la Pensée en Langue Maternelle et/ou Etrangère.). UNESCO report, Paris, 110 pages
- Zock, M., Ferret, O. and D. Schwab, (2010). Deliberate word access : an intuition, a roadmap and some preliminary empirical results. In A. Neustein (éd.) 'International Journal of Speech Technology', 13(4), pp. 107-117, 2010. Springer Verlag
- Zock, M., Sabatier, P. and L. Jakubiec. (2008). Message Composition Based on Concepts and Goals. *International Journal of Speech Technology*, 11(3-4):181–193.