Direkt Profil: A System for Evaluating Texts of Second Language Learners of French Based on Developmental Sequences

Jonas Granfeldt[1] **Pierre Nugues**[2] **Emil Persson**[1] **Lisa Persson**[2]
Fabian Kostadinov[3] **Malin Ågren**[1] **Suzanne Schlyter**[1]

[1]Dept. of Romance Languages
Lund University
Box 201, 221 00 Lund, Sweden

[2]Dept. of Computer Science
Lund University
Box 118, 221 00 Lund, Sweden

[3]Dept. of Computer Science
University of Zurich
CH-8057 Zurich, Switzerland

```
{Jonas.Granfeldt, Malin.Agren, Suzanne.Schlyter}@rom.lu.se
             emil.person@telia.com nossrespasil@hotmail.com
   Pierre.Nugues@cs.lth.se fabian.kostadinov@access.unizh.ch
```

Abstract

Direkt Profil is an automatic analyzer of texts written in French as a second language. Its objective is to produce an evaluation of the developmental stage of students under the form of a grammatical learner profile. *Direkt Profil* carries out a sentence analysis based on developmental sequences, i.e. local morphosyntactic phenomena linked to a development in the acquisition of French.

The paper presents the corpus that we use to develop the system and briefly, the developmental sequences. Then, it describes the annotation that we have defined, the parser, and the user interface. We conclude by the results obtained so far: on the test corpus the systems obtains a recall of 83% and a precision of 83%.

1 Introduction

With few exceptions, systems for evaluating language proficiency and for Computer-Assisted Language Learning (CALL) do not use Natural Language Processing (NLP) techniques. Typically, existing commercial and non-commercial programs apply some sort of pattern-matching techniques to analyze texts. These techniques not only reduce the quality and the nature of the feedback but also limit the range of possible CALL applications.

In this paper, we present a system that implements an automatic analysis of texts freely written by learners. Research on Second Language Acquisition (SLA) has shown that writing your own text in a communicative and meaningful situation with a feedback and/or an evaluation of its quality and its form constitutes an excellent exercise to develop second language skills.

The aim of the program, called *Direkt Profil*, is to evaluate the linguistic level of the learners' texts in the shape of a learner profile. To analyze sentences, the program relies on previous research on second language development in French that itemized a number of specific constructions corresponding to *developmental sequences*.

2 The CEFLE Lund Corpus

For the development and the evaluation of the system, we used the CEFLE corpus (*Corpus Écrit de Français Langue Étrangère de Lund* "Lund Written Corpus of French as a Foreign Language"). This corpus currently contains approximately 100,000 words (Ågren, 2005). The texts are narratives of varied length and levels. We elicited them by asking 85 Swedish high-school students and 22 young French to write stories evoked by a sequence of images. Figure 1 shows pictures corresponding to one of them: *Le voyage en Italie* "The journey to Italy". The goal of the system being to analyze French as a foreign language, we used the texts of the French native speakers as control group.

The following narrative is an example from a be-

ginner learner:

> *Elles sont deux femmes. Elles sont a italie au une vacanse. Mais L'Auto est très petite. Elles va a Italie. Au l'hothel elles demande une chambre. Un homme a le clé. Le chambre est grande avec deux lies. Il fait chaud. C'est noir. Cette deux femmes est a une restaurang. Dans une bar cet deux hommes. Ils amour les femmes. Ils parlons dans la bar. Ils ont tres bien. Le homme et la femme participat a un sightseeing dans la Rome. Ils achetons une robe. La robe est verte. La femme et l'homme reste au un banqe. Ils c'est amour. La femme et l'homme est au une ristorante. es hommes va avec les femmes. L'auto est petite.*

This text contains a certain number of typical constructions for French as a foreign language: parataxis, very simple word order, absence of object pronouns, basic verb forms, agreement errors, spelling mistakes. Research on the acquisition of French as a foreign language has shown that these constructions (and others) appear in a certain systematic fashion according to the proficiency level of the learners. With *Direkt Profil*, we aim at detecting automatically these structures and gathering them so that they represent a grammatical learner profile. This learner profile can ultimately be used to assess learners' written production in French.

3 *Direkt Profil* and Previous Work

Direkt Profil is an analyzer of texts written in French as a foreign language. It is based on the linguistic constructions that are specific to developmental sequences. We created an annotation scheme to mark up these constructions and we used it to describe them systematically and detect them automatically. The analyzer parses the text of a learner, annotates the constructions, and counts the number of occurrences of each phenomenon. The result is a text profile based on these criteria and, possibly, an indication of the level of the text. A graphical user interface (GUI) shows the results to the user and visualizes by different colors the detected structures. It is important to stress that *Direkt Profil* is not a grammar checker.

The majority of the tools in the field can be described as writing assistants. They identify and sometimes correct spelling mistakes and grammatical errors. The line of programs leading to *PLNLP* (Jensen et al., 1993) and NLPWin (Heidorn, 2000) is one of the most notable achievements. The grammatical checker of *PLNLP* carries out a complete parse. It uses binary phrase-structure rules and takes into account some dependency relations. *PLNLP* is targeted primarily, but not exclusively, to users writing in their mother tongue. It was created for English and then applied to other languages, including French.

Other systems such as *FreeText* (Granger et al., 2001) and *Granska* (Bigert et al., 2005) are relevant to the CALL domain. *FreeText* is specifically designed to teach language and adopts a interactive approach. It uses phrase-structure rules for French. In case of parsing failure, it uses relaxed constraints to diagnose an error (agreement errors, for example). *Granska*, unlike *FreeText*, carries out a partial parsing. The authors justify this type of analysis by a robustness, which they consider superior and which makes it possible to accept more easily incorrect sentences.

4 An Analysis Based on Developmental Sequences

The current systems differ with regard to the type of analysis they carry out: complete or partial. The complete analysis of sentences and the correction of errors are difficult to apply to texts of learners with (very) low linguistic level since the number of unknown words and incorrect sentences are often extremely high.

We used a test corpus of 6,842 words to evaluate their counts. In the texts produced by learners at the lowest stage of development, Stage 1, nearly 100% of the sentences contained a grammatical error (98.9% were incorrect[1]) and 24.7% of the words were unknown.[2] At this stage of development, any complete analysis of the sentences seems very difficult to us. On the other hand, in the control group the

---

[1] An "incorrect sentence" was defined as a sentence containing at least one spelling, syntactic, morphological, or semantic error.

[2] An "unknown word" is a token that does not appear in the lexicon employed by the system (ABU CNAM, see below)
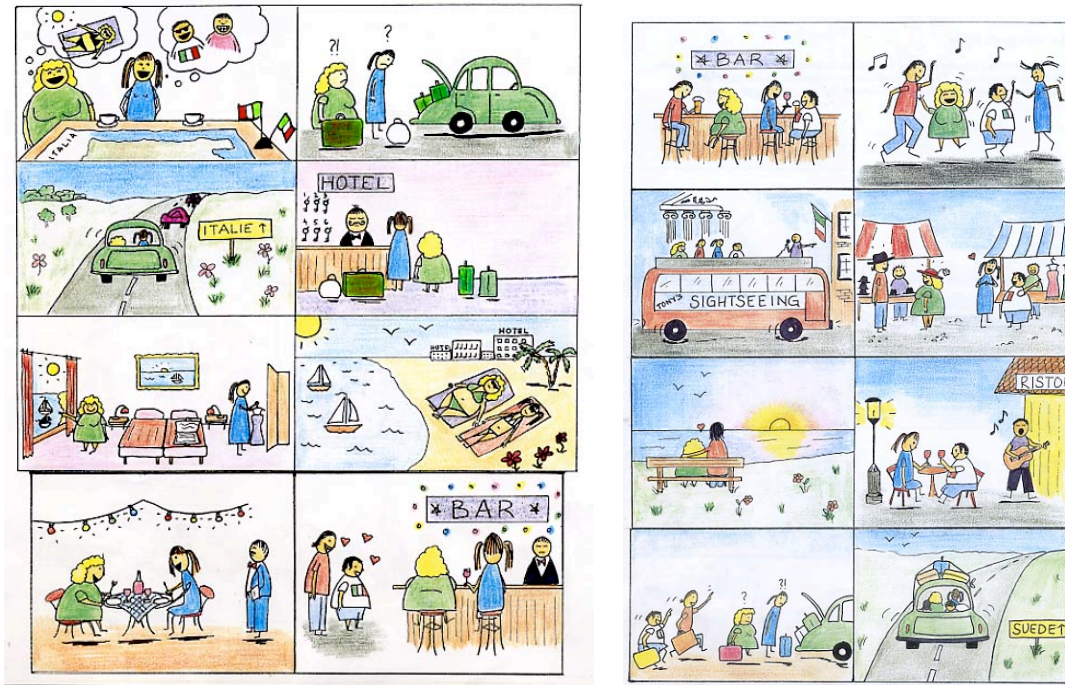
Figure 1: *Le voyage en Italie* "The journey to Italy".

corresponding figures are 32.7% for incorrect sentences and 10.6% for unknown words. More importantly, this analysis shows that using a quantification of "unknown words" and "incorrect sentences" only is insufficient to define the linguistic level of learners' texts. Learners at Stage 3 have in fact fewer incorrect sentences than learners from Stage 4 (70.5% vs. 80.2%). Moreover, the percentage of unknown words in the control group (the natives) is slightly higher than that of learners from the Stage 4 (10.6% vs. 10.4%). Thus, the simple count of errors is also insufficient to distinguish more advanced learners from natives. To identify properly and to define learners of various linguistic levels, we need more detailed analyses and more fine-grained measures. This is exactly the purpose of the developmental sequences and learner profiles implemented in *Direkt Profil*.

5   Developmental Sequences in French

Direkt Profil carries out an analysis of local phenomena related to a development in the acquisition of French. These phenomena are described under the form of developmental sequences. The sequences are the result of empirical observations stemming from large learner corpora of spoken language (Bartning and Schlyter, 2004). They show that certain grammatical constructions are acquired and can be produced in spontaneous spoken language in a fixed order. Clahsen and al. (1983) as well as Pienemann and Johnston, (1987) determined developmental sequences for German and spoken English. For spoken French, Schlyter (2003) and Bartning and Schlyter (2004) proposed 6 stages of development and developmental sequences covering more than 20 local phenomena. These morphosyntactic phenomena are described under the form of local structures inside the verbal or nominal domain. Table 1 shows a subset of these phenomena. It is a matter of current debate in field of SLA to what extent these developmental sequences are independent of the mother tongue.

The horizontal axis indicates the temporal development for a particular phenomenon: The developmental sequence. The vertical axis indicates the set of grammatical phenomena gathered in such way that they make up a "profile" or a stage of acquisition. To illustrate better how this works, we will compare the C (finite verb forms in finite contexts) and G (object pronouns) phenomena.

55

At Stage 1, the finite and infinitive forms coexist in finite contexts. As the main verb of the sentence, we find in the learners' production *je parle* (transcription of /je parl/ analyzed as a "finite form") as well as /je parle/ i.e. **je parler* or **je parlé*. The current estimation is that in Stage 1, there are between 50 and 75% of finite forms in finite contexts. At Stage 4, the percentage of finite forms has increased to 90–98%. For this morphological phenomenon, the developmental sequence describes a successive "morphologization".

The G phenomenon concerns the developmental sequence of object pronouns. The first object pronouns are placed in a postverbal position according to the scheme Subject-Verb-Object (SVO), e.g. **je vois le/la/lui* (instead of *je le/la vois*). At Stage 3, learners can produce phrases according to the SvOV scheme (Pronoun-Auxiliary-Object-Verb): *Je veux le voir* (correct) but also **j'ai le vu* (incorrect). At Stage 5, we observe *je l'ai vu*. For this syntactic phenomenon, the developmental sequence describes a change in the linear organization of the constituents.

6 Annotation

The concept of group, either noun group or verb group, correct or not, represents the essential grammatical support of our annotation. The majority of syntactic annotation standards for French takes such groups into account in one way or another. Gendner et al. (2004) is an example that reconciles a great number of annotations. These standards are however insufficient to mark up all the constructions in Table 1.

We defined a text annotation specific to *Direkt Profil* based on the inventory of the linguistic phenomena described by Bartning and Schlyter (2004) (Table 1). We represented these phenomena by decision trees whose final nodes correspond to a category of analysis.

The annotation uses the XML format and annotates the texts using 4 layers. Only the 3rd layer is really grammatical:

- The first layer corresponds to the segmentation of the text in words.

- The second layer annotates prefabricated expressions or sentences (e.g. *je m'appelle*).

These structures correspond to linguistic expressions learned "by heart" in a holistic fashion. It has been shown that they have a great importance in the first years of learning French.

- The third layer corresponds to a chunk annotation of the text, restricted to the phenomena to identify. This layer marks up simultaneously each word with its part-of-speech and the verb and noun groups to which they belong. The verb group incorporates subject clitic pronouns. The XML element `span` marks the groups and features an attribute to indicate their class in the table. The `tag` element annotates the words with attributes to indicate the lemma, the part-of-speech, and the grammatical features. The verb group in the sentence *Ils parlons dans la bar* extracted from the learner text above is annotated as:
```
<span class="p1_t1_c5131"><tag
pos="pro:nom:pl:p3:mas">Ils</tag>
<tag pos="ver:impre:pl:p1">
parlons </tag></span> dans la
bar.
```
The class denoted `p1_t1_c5131` corresponds to a "finite lexical verb, no agreement".

- The fourth layer counts structures typical of an acquisition stage. It uses the `counter` XML element, `<counter id="counter.2" counter_name="passe_compose" rule_id="participe_4b" value="1"/>`.

7 Implementation

The running version of *Direkt Profil* is restricted to the analysis of the verb groups and clitic pronouns. For each category in Table 1, the program identifies the corresponding constructions in a text and counts them.

The analyzer uses manually written rules and a lexicon of inflected terms. The variety of the constructions contained in the corpus is large and in order not to multiply the number of rules, we chose a constraint reinforcement approach. Conceptually, the analyzer seeks classes of phrase structures in which all the features are removed. It gradually identifies the structures while varying the feature

Ph.	**Stages**	1	2	3	4	5	6
A.	% of sentences containing a verb (in a conversation)	20–40%	30–40%	50%	60%	70%	75%
B.	% of lexical verbs showing +/-finite opposition (types)	No opp.; % in finite forms 1–3sg	10–20% of types in opposition	About 50% in opposition	Most in opposition	All in opposition	+
C.	% of finite forms of lexical verbs in obligatory contexts (occurrences)	Finite forms 50%–75%	Finite forms 70–80%	Finite forms: 80–90%	Finite forms: 90–98%	Finite forms: 100%	+
D.	1st, 2nd, 3rd pers. sing. (copula/aux) *est, a, va*	No opposition: *J'ai/ c'est*	Opposition *j'ai – il a je suis – il est*	Isolated errors *je va*, *je a*	+	+	+
E.	% of 1st pers. plural S-V agreement *nous* V-*ons* (occurrences)	–	70–80%	80–95%	Errors in complex constructions	+	+
F.	3rd pers. plural S-V agreement with *viennent, veulent, prennent*	–	– *ils *prend*	Isolated cases of agreement	50% of cases with agreement	Some problems remain	+
G.	Object pronouns (placement)	–	SVO	S(v)oV	SovV appears	Productive	+ (*y, en*)
H.	% of gender agreement Article-Noun (occurrences)	55–75%	60–80%	65–85%	70–90%	75–95%	90–100%

Table 1: Developmental sequences adapted from Schlyter (2003); Bartning and Schlyter (2004).
Legend: – = no occurrences; + = acquired at a native-like level; aux = auxiliary; pers. = person; S-V = Subject-Verb

values. The recognition of the group boundaries is done by a set of closed-class words and heuristics inside the rules. It thus follows an old but robust strategy used in particular by Vergne (1999), *inter alia*, for French.

Direkt Profil applies a cascade of three sets of rules to produce the four annotation layers. The first unit segments the text in words. An intermediate unit identifies the prefabricated expressions. The third unit annotates simultaneously the parts-of-speech and the groups. Finally, the engine creates a group of results and connects them to a profile. It should be noted that the engine neither annotates all the words, nor all segments. It considers only those which are relevant for the determination of the stage. The engine applies the rules from left to right then from right to left to solve certain problems of agreement.

The rules represent partial structures and are divided into a condition part and an action part. The condition part contains the search parameters. It can be a lemma, a regular expression, or a class of inflection. The engine goes through the text and applies the rules using a decision tree. It tests the condition part to identify the sequences of contiguous words. Each rule produces a positive ("match") or negative ("no match") result. The rules are applied according to the result of the condition part and annotate the text, count the number of occurrences of the phenomenon, and connect to another rule. By traversing the nodes of the tree, the engine memorizes the rules it has passed as well as the results of the condition parts of these rules. When arriving at a final node, the engine applies the action parts of all the rules.

The engine finds the words in a dictionary of inflected terms. It does not correct the spelling mistakes except for the accents and certain stems. Learners frequently build erroneous past participles inferring a wrong generalization of stems. An example is the word *prendu (taken) formed on the stem *prend|re* and of the suffix *-u*.

We used a lexicon available from the Association des Bibliophiles Universels' web site (http://abu.cnam.fr/) that we corrected and transposed into XML. We also enriched it with verb stems.

8   Interface

Direkt Profil merges the annotation levels in a result object. This object represents the original text, the annotation, the trace of the rule application, and the counters. The result object, which can be saved, is then transformed by the program to be presented to the user. The display uses the XHTML 1.1 specifications which can be read by any Web browser. *Direkt Profil* has a client-server architecture where the server carries out the annotation of a text and the client collects the text with an input form and interacts with the user.

Figure 2 shows a screenshot of *Direkt Profil*'s GUI displaying the analysis of the learner text above. The interface indicates to the user by different colors all the structures that the analyzer detected.

9   Results and Evaluation

We evaluated *Direkt Profil* with a subset of the CEFLE corpus. We chose 20 texts randomly distributed on 4 learner stages. We also used 5 texts coming from the control group. In this version, we did not test the correction of the misspelled words: accent and stems. Table 2 shows some statistics on the size of the texts and Table 3 shows the results in the form of recall and precision.

The results show that *Direkt Profil* detects well the desired phenomena. It reveals also interesting differences according to the levels of the texts. The results show that *Direkt Profil* analyzes better the learner texts than the texts from the native French adolescents (control group). Without knowing exactly why, we note that it suggests that the adopted strategy, which aims at analyzing texts in French as a foreign language, seems promising.

10   Conclusion and Future Work

We presented a system carrying out a machine analysis of texts based on developmental sequences. The goal is to produce a learner profile. We built a parser and developed a set of rules to annotate the texts. *Direkt Profil* is integrated in a client-server architecture and has an interface allowing the interaction with the user.

The results show that it is possible to describe the vast majority of the local structures defined by the
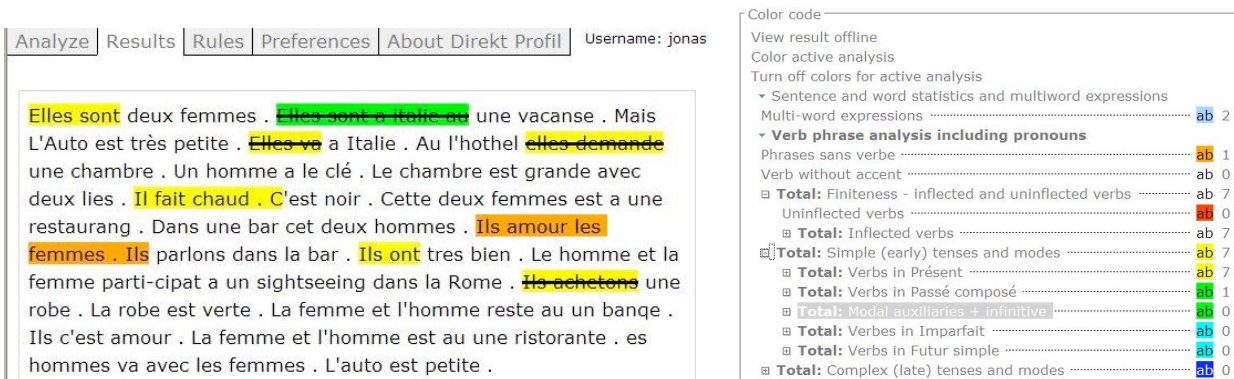
Figure 2: The graphical user interface.

	Stage 1	Stage 2	Stage 3	Stage 4	Control	Total
Number of analyzed texts	5	5	5	5	5	25
Word count	740	1233	1571	1672	1626	6842
Sentence count	85	155	166	126	107	639
Average text length (in words)	148	247	314	334	325	274
Average length of sentences (in words)	8.7	7.9	9.5	13.3	15.2	10.9

Table 2: Test corpus.

	Stage 1	Stage 2	Stage 3	Stage 4	Control	Total
Reference structures	23	97	101	119	85	425
Detected structures	27	98	100	112	92	429
Correctly detected structures	15	81	89	96	73	354
Non detected structures	5	16	12	20	11	()64
Overdetected structures	10	17	11	17	19	()74
Recall	65%	84%	88%	81%	86%	83%
Precision	56%	83%	89%	86%	79%	83%
F-measure	0.6	0.83	0.89	0.83	0.82	0.83

Table 3: Results.

developmental sequences under the form of rules. *Direkt Profil* can then detect them and automatically analyze them. We can thus check the validity of the acquisition criteria.

In the future, we intend to test *Direkt Profil* in teaching contexts to analyze and specify, in an automatic way, the grammatical level of a learner. The program could be used by teachers to assess student texts as well as by the students themselves as a self-assessment and as a part of their learning process.

A preliminary version of *Direkt Profil* is available on line from this address `http://www.rom.lu.se:8080/profil`

References

Malin Ågren. 2005. Le marquage morphologique du nombre dans la phrase nominale. une étude sur l'acquisition du français L2 écrit. Technical report, Institut d'études romanes de Lund. Lund University.

Inge Bartning and Suzanne Schlyter. 2004. Stades et itinéraires acquisitionnels des apprenants suédophones en français l2. *Journal of French Language Studies*, 14(3):281–299.

Johnny Bigert, Viggo Kann, Ola Knutsson, and Jonas Sjöbergh. 2005. Grammar checking for Swedish second language learners. In *CALL for the Nordic Languages*, Copenhagen Studies in Language, pages 33–47. Copenhagen Business School, Samfundslitteratur.

Harald Clahsen, Jürgen M. Meisel, and Manfred Pienemann. 1983. *Deutsch als Fremdsprache. Der Spracherwerb ausländischer Arbeiter*. Narr, Tübingen.

Véronique Gendner, Anne Vilnat, Laura Monceaux, Patrick Paroubek, and Isabelle Robba. 2004. Les annotations syntaxiques de référence peas. Technical report, LIMSI, Orsay. http://www.limsi.fr/Recherche/CORVAL/easy/PEAS_reference_annotations_v1.6.html.

Sylviane Granger, Anne Vandeventer, and Marie-Josée Hamel. 2001. Analyse de corpus d'apprenants pour l'ELAO basé sur le TAL. *Traitement automatique des langues*, 42(2):609–621.

George E. Heidorn. 2000. Intelligent writing assistance. In Robert Dale, Hermann Moisl, and Harold Somers, editors, *Handbook of Natural Language Processing*. Marcel Dekker.

Karen Jensen, George E. Heidorn, and Stephen D. Richardson. 1993. *Natural Language Processing: The PLNLP Approach*. Kluwer Academic Publishers.

Manfred Pienemann and Malcolm Johnston. 1987. Factors influencing the development of second language proficiency. In David Nunan, editor, *Applying second language acquisition research*, pages 45–141. National Curriculum Resource Centre, Adelaide.

Suzanne Schlyter. 2003. Stades de développement en français L2. Technical report, Institut d'études romanes de Lund, Lund University. http://www.rom.lu.se/durs/STADES_DE_DEVELOPPEMENT_EN_FRANCAIS_L2.PDF.

Jacques Vergne. 1999. *Étude et modélisation de la syntaxe des langues à l'aide de l'ordinateur. Analyse syntaxique automatique non combinatoire. Synthèse et Résultats*. Habilitation à diriger des recherches, Université de Caen, 29 septembre.