

**DE L'ACQUISITION DES LANGUES  
À L'ÉVALUATION EN FLE – LE LOGICIEL  
DIREKT PROFIL EN APPLICATION**

**Jonas Granfeldt & Malin Ågren**

Centre de langues et de littérature, Université de Lund, Suède

**Mots-clés**

Français langue étrangère – L1 suédois – itinéraires d'acquisition – *Direkt Profil* – évaluation

**Keywords**

French as foreign language – L1 Swedish – acquisitional orders – *Direkt Profil* – assessment

**Résumé**

*Direkt Profil* est un analyseur morphosyntaxique de français langue étrangère, basé sur un modèle de stades de développement grammatical des apprenants suédophones de français. *Direkt Profil* produit automatiquement un profil linguistique à partir d'un texte d'apprenant et d'une estimation de son niveau de développement morphosyntaxique. L'article avance l'idée selon laquelle un logiciel de ce type pourrait servir d'interface entre les recherches en acquisition et la didactique des langues, notamment dans l'évaluation diagnostique et formative.

**Abstract**

*Direkt Profil* is a morphosyntactic analyzer for French as a foreign language, based on a model of grammatical stages of development in Swedish learners of French. *Direkt Profil* automatically produces a linguistic profile of a learner text and an estimation of its stage of development. The paper discusses the idea that software of this type could function as an interface between research on language acquisition and language didactics, especially in the domains of diagnostic and formative assessment.

## Introduction

Durant ces dernières décennies, les recherches sur l'acquisition des langues étrangères (RAL) ont permis non seulement de décrire les étapes du développement linguistique de l'apprenant en langue étrangère, mais aussi d'affiner la compréhension des facteurs et des processus qui influencent l'appropriation d'une langue étrangère, notamment le rôle de la langue première (L1). Or, malgré l'importance potentielle de ces acquis pour les recherches en didactique des langues (DDL), plusieurs auteurs constatent un manque de dialogue entre les RAL et la DDL (Coste, 1992 ; Véronique, 2000 ; Devitt, 2004, Trévisiol-Okamura & Komur-Thillo, 2011). Les raisons pour cette quasi-absence d'échanges varient selon les auteurs, mais ce qui semble la sous-tendre est l'idée d'une différence fondamentale entre ces deux projets scientifiques :

*Les recherches sur l'acquisition relèvent des sciences du langage, discipline qui vise à dégager des régularités et des lois du fonctionnement linguistique et communicatif, tandis que la didactique des langues est une tekhnè, une discipline de l'action et de l'intervention dans le domaine de l'enseignement-apprentissage linguistique. (Véronique, 2009 : 326).*

Au cœur de ce constat, nous apercevons l'idée selon laquelle les acquis des RAL représentent *un savoir*, aspirant à une certaine décontextualisation pour permettre la généralisation à travers le temps quels que soient les situations et les individus (cf. l'accent sur *régularités, lois*), alors que la didactique représente *un savoir-faire*, nécessairement contextualisé et situé pour capter les spécificités propres à chaque situation didactique (cf. *action, intervention*).

Formulé ainsi, il est possible que le dialogue limité entre ces deux disciplines soit, du moins en partie, dû à l'absence relative d'*interfaces* adaptées qui permettraient de dépasser les limites de chacune des deux disciplines et de leurs objectifs scientifiques, une interface dont les qualités essentielles seraient celles de la systématité, de l'ouverture et de l'adaptation. Selon une définition possible, une *interface* est la couche limite entre deux éléments au travers de laquelle ont lieu des échanges et des interactions.<sup>1</sup> Dans cet article, nous avancerons l'hypothèse générale selon laquelle les Technologies d'Interaction et Communication (TIC), couplés au Traitement Automatique de Langues (TAL), sous forme d'un système expert doté des acquis des recherches en RAL, permettent aux chercheurs en DDL et aux enseignants de langues étrangères d'accéder aux acquis des RAL, mais aussi de contribuer au progrès de ces mêmes recherches, notamment dans les domaines de l'évaluation diagnostique et formative.

L'objectif du présent article est de discuter comment le développement d'un logiciel pour l'analyse automatique du français langue étrangère (FLE), *Direkt Profil*

---

1. Il s'agit là d'une définition générale provenant de Wikipedia.org. Le terme interface est avant tout utilisé dans le domaine informatique, mais ici nous envisageons le terme dans son sens large et abstrait.

(Granfeldt et al., 2005 ; Granfeldt & Nugues, 2007) a déjà permis de relier ces deux domaines dans une étude exploratoire (Granfeldt & Ågren, 2008, sous presse). *Direkt Profil* a été construit sur la base de résultats de plusieurs décennies de recherches sur l'acquisition du FLE. La conception du logiciel se base sur un modèle d'itinéraires d'acquisition et de stades de développement traçant les progrès successifs de l'apprenant débutant au locuteur très avancé de FLE. Le modèle a pris son départ dans le développement de la grammaire de l'apprenant suédophone du français parlé (Bartning & Schlyter, 2004), mais il a ultérieurement été adapté au développement du FLE écrit (Ågren, 2008) et appliqué aux apprenants d'autres L1 germaniques (Housen, Kemps & Pierrard, 2009).

Dans un premier temps, nous donnerons un bref aperçu du modèle de Bartning et Schlyter (2004), pour ensuite présenter un exemple typique du développement des apprenants suédophones dans le domaine du FLE écrit : l'accord sujet-verbe en nombre. Ceci nous permettra de discuter une influence possible de la L1. La section 3 sera consacrée à une brève présentation du logiciel *Direkt Profil*, son fonctionnement et son architecture. La dernière section discutera le rôle de *Direkt Profil* en tant qu'interface entre les RAL et un aspect de la DDL, à savoir l'évaluation des compétences linguistiques.

## 1. Développement des compétences morphosyntaxiques en FLE

Les recherches en acquisition du FLE ont connu une productivité remarquable en Suède pendant les dernières décennies. En 2004, Bartning et Schlyter ont publié une synthèse de ces travaux basés sur les deux grands corpus oraux 'Corpus Lund' et 'Corpus Interfra'<sup>2</sup>. S'appuyant sur des itinéraires acquisitionnels observés pour une vingtaine de phénomènes morphosyntaxiques différents (p.ex. la finitude, l'accord sujet-verbe, le temps et l'aspect, le placement de la négation, la subordination, etc.), ces chercheuses ont pu observer un développement relativement stable et semblable chez différents apprenants, guidés et non-guidés, d'âges variés et exposés à différents types d'input français. L'idée principale exprimée par Bartning et Schlyter (2004) est que les apprenants suédophones de FLE semblent passer par les mêmes étapes d'acquisition pour aller vers la langue cible. Il est donc possible de décrire les caractéristiques de l'interlangue d'un apprenant *A* à un moment *T* de son développement sous forme de *profil morphosyntaxique*.

Bartning et Schlyter (2004) décrivent ces observations sous forme d'un modèle basé sur six stades de développement allant des débutants aux apprenants très avancés. Le Tableau 1 illustre le fait que les itinéraires d'acquisition de phénomènes spécifiques (sur l'axe horizontal) sont liés à un nombre limité de stades de

---

2. Pour toute information, voir Bartning et Schlyter (2004) ; pour le Corpus Interfra : <http://www.fraitaklass.su.se/english/interfra/home-interfra>.

développement (sur l'axe vertical).<sup>3</sup> A titre d'exemple, le profil morphosyntaxique d'un apprenant au stade 3 (intermédiaire) a été mis en valeur dans le Tableau 1.

Stade de développement	Stade 1 Initial	Stade 2 Post-initial	Stade 3 Intermédiaire	Stade 4 Avancé bas	Stade 5 Avancé moyen	Stade 6 Avancé supérieur
% Finitude verbes lexicaux	50–75 %	70–80 %	80–90 %	90–98 %	100 %	100 %
% accord sujet-verbe (1 <sup>e</sup> pl)	–	70–80 %	80–95 %	100 %	100 %	100 %
% Accord sujet-verbe (3 <sup>e</sup> pl) Verbes irréguliers lexicaux	–	–	Quelques occurrences	≈ 50 %	Peu d'erreurs	100 %
Temps Mode Aspect	Présent	Présent (Futur.pr) (PC)	Présent Futur.pr PC (Impf)	Présent Futur.pr PC Impf	Présent Futur.pr PC Impf Pluperf. Futur.sim Cond.	Présent Futur.pr PC Impf PQP. Futur.sim Cond. Subj.
Placement pronom objet	–	SVO	S(v)oV	SovV emerge	SovV productif	SovV acquis (+ en / y)
% Accord en genre	55–75 %	60–80 %	65–85 %	70–90 %	75–95 %	90–100 %

Tableau 1. Séquences acquisitionnelles et stades de développement en FLE.

Selon les auteures, ce modèle reflète l'acquisition du français de l'apprenant suédophone dans des situations orales spontanées où il doit avoir recours à ses connaissances implicites et automatisées. Les auteures attribuent une certaine importance à l'influence de la L1 sur le développement observé, surtout dans le domaine syntaxique (en ce qui concerne l'ordre des mots), mais elles prétendent que le modèle est généralisable aux apprenants d'autres langues scandinaves et germaniques, ce qui a récemment été vérifié par Housen, Kemps et Pierrard (2007, 2009) pour des apprenants flamands de FLE.

3. Pour une version plus élaborée du Tableau 1, voir Bartning et Schlyter (2004 : 294).

## 2. Développement des compétences morphosyntaxiques en FLE écrit

En français, l'écrit présente de nombreuses marques morphosyntaxiques que l'on ne distingue pas à l'oral (par exemple accord en nombre, genre, personne). Or, le modèle de Bartning et Schlyter (2004) présenté ci-dessus ne rend pas compte du développement des distinctions morphologiques silencieuses du français écrit. Pour cette raison, Ågren (2008) a voulu tester l'hypothèse suivant laquelle la morphologie du français écrit et plus précisément celle du nombre évolue selon un itinéraire systématique chez les apprenants suédophones de français à l'écrit. Lors du passage de l'oral à l'écrit, il faut noter l'accroissement de la divergence entre les langues source (suédois) et cible (français) des apprenants. La correspondance oral-écrit est plus transparente en suédois qu'en français (Fayol & Jaffré, 2008). En ce qui concerne le nombre à l'oral et à l'écrit, le suédois présente un accord à l'intérieur du syntagme nominal alors qu'il y a absence d'accord sujet-verbe en nombre et en personne (p.ex. *jag spelar* 'je joue', *de spelar* 'ils jouent', *vi spelar* 'nous jouons', etc.). La morphologie verbale est donc beaucoup plus riche en français écrit qu'en suédois (parlé et écrit), ce qui ne sera pas sans importance pour le développement observé dans ce domaine (voir ci-dessous).

Dans le but de tracer le développement morphosyntaxique du FLE à l'écrit, Ågren a créé une base de données de textes intitulée CEFLE<sup>4</sup>. Les tâches utilisées dans ce corpus correspondent à des récits de souvenirs personnels ou de séries d'images, tâches qui mettent l'accent sur la transmission d'un message et non pas sur la forme et qui invitent à la mise en oeuvre de connaissances automatisées en FLE. Ce corpus inclut près de 400 textes écrits en français L2 par une centaine d'apprenants au lycée en Suède qui appartiennent aux stades initial (1), post-initial (2), intermédiaire (3) et avancé bas (4) selon le modèle de Bartning et Schlyter (2004), ainsi que par un groupe contrôle de scripteurs natifs<sup>5</sup>.

L'analyse de la production écrite en FLE de Ågren (2008) montre un développement progressif et rapide de la morphologie silencieuse du nombre entre les quatre stades observés. Le schéma 1, qui résume le résultat de l'étude dans le domaine de l'accord sujet-verbe en nombre (tous les groupes de verbes confondus), illustre des différences significatives entre tous les stades de développement (p.<.05), sauf entre le stade avancé bas et les adolescents natifs du groupe contrôle. Il est donc à noter que, déjà à un stade avancé bas, les apprenants L2 atteignent le niveau du groupe contrôle dans ce domaine, ce qui n'est pas le cas pour beaucoup d'autres phénomènes morphosyntaxiques (voir stade 4, Tableau 1).

4. Le Corpus Écrit de Français Langue Étrangère (CEFLE) est en libre accès sur Internet : <http://project.sol.lu.se/cefle>.

5. CEFLE est basé sur des textes en FLE écrits dans un contexte scolaire (au lycée en Suède). Ces textes n'atteignent pas (encore) les deux stades les plus avancés du modèle de Bartning et Schlyter (2004). Pour cette raison, le corpus n'inclut pas de textes appartenant aux stades 5 (avancé moyen) et 6 (avancé supérieur).

De manière générale, on ne trouve pas d'apprenants suédophones de FLE en phase initiale chez qui la morphologie silencieuse du nombre est complètement absente (cf. la production écrite des enfants francophones, Fayol, 2003). Au stade initial, la distinction du pluriel est produite surtout là où elle est sémantiquement motivée, c'est-à-dire sur le quantifieur et sur le nom/pronom. Or simultanément l'accord sujet-verbe en nombre est la plupart du temps omis, comme dans l'exemple suivant.

(1) Apprenant de FLE : Stade initial (exemple tiré de Ågren, 2008 : 176)

*Puis elles \*dîner a la restaurant. Elles \*boisse un carafe de la vin. Deux garçons \*aime les filles. Ils \*parler et \*boisse biere et vin. Apres ils ecoutent la musique.*

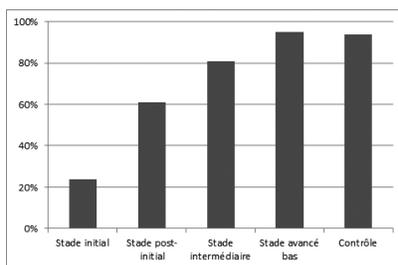


Schéma 1. Développement de l'accord sujet-verbe en nombre en FLE écrit.

Au stade post-initial déjà intervient une phase très créative où les apprenants essaient de mettre en œuvre le système d'accord verbal du français écrit, sans pour autant l'avoir entièrement automatisé. Aux stades post-initial et intermédiaire, les textes font preuve d'une variation de formes qui témoigne d'un système interlangagier en développement. L'exemple (2) illustre des formes dites 'créatives', où le graphème *-nt* est employé avec différents types de radicaux, un procédé qui apparaît en abondance à côté de formes correctement accordées et d'omissions persistantes.

(2) Accord sujet-verbe en nombre produit aux stades post-initial et intermédiaire (exemples tirés de Ågren, 2008 : 176-179)

*Après ils ecoutent la musique.*

*Elles sont sur le plage.*

*Le soir elles \*allent à une restaurant. (cf. vont)*

*Le jour après ils \*faitent le voyage. (cf. font)*

*Les filles \*dirent au revoir. (cf. disent)*

Finalement, au stade avancé bas, les apprenants de FLE atteignent un taux d'accord sujet-verbe en nombre semblable à celui du groupe contrôle francophone. Les formes créatives ont presque disparu et le taux d'accord dépasse 90 % des contextes obligatoires. Cette évolution graduelle s'oppose nettement à celle

observée par Granget (2006) dans des textes écrits en FLE par des apprenants germanophones, qui sont très performants dans ce domaine morphosyntaxique dès le stade initial. Ce résultat pourrait être le signe d'une influence de la L1 des apprenants sur leur production écrite en FLE. Contrairement aux apprenants germanophones, les apprenants suédophones ont une L1 qui ne présente pas d'accord sujet-verbe, ce qui rendrait plus laborieuse l'acquisition de cette dimension morphologique du français écrit.

Dans la section suivante, nous allons décrire comment les résultats présentés dans les sections 1 et 2 ont servi de base pour l'élaboration d'un logiciel pour l'analyse automatique du profil grammatical d'un texte écrit en FLE.

### 3. Direkt Profil – un analyseur morphosyntaxique automatique du FLE

*Direkt Profil* est un analyseur morphosyntaxique conçu pour analyser des textes de FLE et construit à l'aide de techniques de TAL robustes, c'est-à-dire des techniques dont la maturité dans la discipline est telle que leur résultat est stable et fiable. Son but est double : 1) produire automatiquement un profil grammatical des constructions, correctes ou non, utilisées dans un texte d'apprenant, et 2) lui attribuer le stade de développement reflété dans le texte. Le système, sa construction et ses performances ont fait l'objet de présentations et d'évaluations détaillées dans des articles précédents (cf. Granfeldt, 2006 ; Granfeldt & Nugues, 2007 ; Granfeldt *et al.*, 2005), raison pour laquelle nous nous contenterons ici de donner une brève description de son fonctionnement et de son architecture.

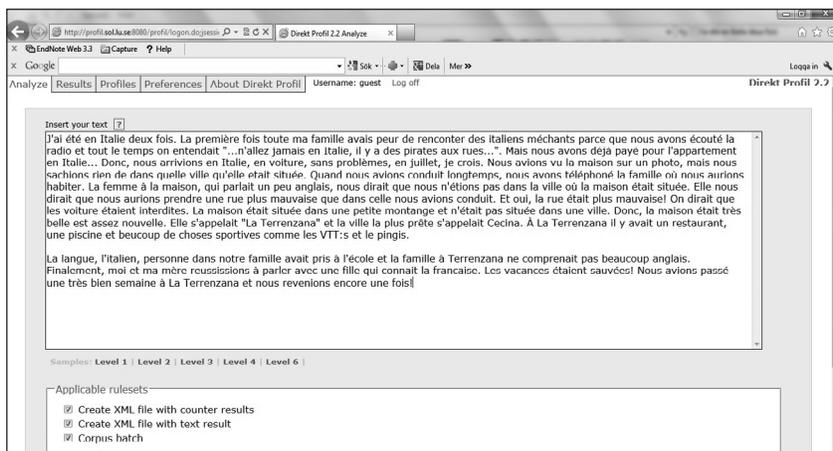
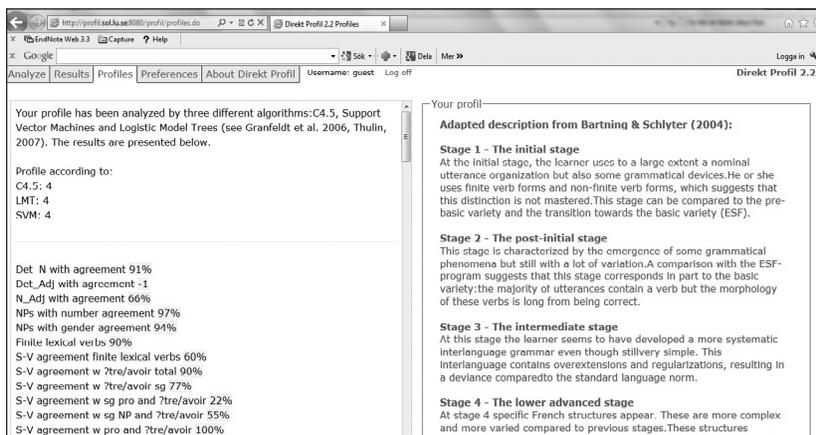


Schéma 2. Capture d'écran d'un texte d'apprenant avant l'analyse de *Direkt Profil*.

Schéma 3. Capture d'écran d'un texte d'apprenant après l'analyse de *Direkt Profil*.

### 3.1. Fonctionnement

*Direkt Profil* fonctionne en mode client-serveur où le serveur réalise l'annotation d'un texte et le client, intégré à un navigateur, prend en charge l'affichage et l'interaction. Aucune installation au préalable n'est nécessaire : il suffit de se connecter sur la page web de *Direkt Profil* ([www.profil.sol.lu.se](http://www.profil.sol.lu.se)) où l'accès au système est libre. L'interface du programme est conçue pour que les chercheurs (ou les enseignants ou les apprenants) puissent copier des textes écrits et les soumettre à l'analyseur. Après la connexion, l'utilisateur se trouve face à un écran d'analyse où il écrit ou colle un texte (schéma 2). L'utilisateur appuie sur le bouton « Analyser le texte » et le programme identifie certaines structures caractéristiques du développement morphosyntaxique en FLE (voir sections 1 et 2 ci-dessus). Après quelques secondes, le système affiche les résultats sur l'écran (schéma 3).

L'écran indiquant les résultats (schéma 3) est divisé en deux parties. Dans la partie de gauche, les structures détectées sont colorées selon des codes que nous avons conçus au préalable. La partie de droite énumère les catégories d'analyse et le nombre d'occurrences de chaque catégorie. Grâce à un script Java, il est possible de cliquer sur le nombre d'occurrences pour faire apparaître une liste contenant les structures correspondantes dans le texte analysé. Dans un dernier temps (schéma 4), le système offre une estimation du stade de développement (tel qu'il se manifeste dans le texte), accompagné d'une brève présentation des caractéristiques de chaque stade selon le modèle de Bartning et Schlyter (2004).

### 3.2. Architecture

*Direkt Profil* se compose d'un analyseur noun morphosyntaxique qui est relié à un module d'apprentissage automatique (voir schéma 5). D'abord, nous avons construit une annotation en XML de la plupart des phénomènes morphosyntaxiques du modèle (voir Tableau 1), mais aussi d'un certain nombre d'indices quantitatifs (longueur moyenne des phrases, etc.) et lexicaux (diversité lexicale, fréquence des mots utilisés).<sup>6</sup>

Schéma 4. Capture d'écran de l'indication du stade de développement d'un texte analysé (à gauche) et la description des stades (à droite).

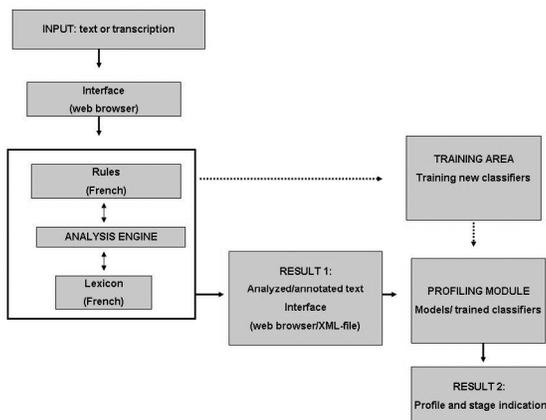


Schéma 5. Architecture schématique de *Direkt Profil*.

6. Pour une discussion de la pertinence et du poids respectif de ces indices pour rendre compte du développement linguistique en FLE, voir Granfeldt (2006) et Granfeldt & Nugues (2007).

De façon conceptuelle, l'analyseur (voir schéma 5 *Analysis engine*) recherche dans un premier temps des groupes syntaxiques (groupes nominaux (GN) et groupes verbaux (GV)) sans considérer les traits grammaticaux des constituants individuels qui les composent. Il identifie ensuite progressivement les constituants des GN et GV en tentant d'affecter des valeurs à chaque constituant. *Direkt Profil* applique en cascade trois ensembles de règles pour produire quatre niveaux d'annotation. Le premier ensemble segmente le texte en mots. Un deuxième ensemble identifie les expressions figées. Le troisième ensemble annote simultanément les parties du discours, les GN et les GV. La reconnaissance des limites des GV et GN se fait par un ensemble de mots vides ou mots grammaticaux (Vergne, 1999). Finalement, le moteur crée un groupe de résultats relié au stade de développement de l'apprenant.

Il est à noter que le moteur n'annote pas tous les mots, ni tous les segments. Il ne considère que ceux qui sont pertinents par rapport aux constructions du modèle de Bartning et Schlyter (2004) et, ultérieurement, pour l'attribution du stade de développement du texte de l'apprenant. Le moteur applique les règles de gauche à droite puis de droite à gauche pour résoudre certains problèmes d'accord.

L'indication des stades de développement est possible grâce au module d'apprentissage automatique (voir schéma 5. *Training area*). Nous avons entraîné trois algorithmes différents sur un corpus annoté manuellement (le corpus CEFLE, voir ci-dessus) pour produire ce qu'on appelle des classificateurs (il s'agit de vecteurs d'indices linguistiques provenant des textes d'apprenants, tels que p. ex. l'accord en genre ou en personne, l'emploi des temps verbaux ou du subjonctif, etc.). Dans la version courante, les algorithmes sont entraînés sur les récits de ce corpus, mais grâce à la modularité du système (voir schéma 5) de nouveaux classificateurs peuvent être établis si besoin est (ce point est discuté ci-dessous). Les résultats actuels montrent que pour un classement en trois stades, le taux de convergence entre être humain et machine se trouve autour de 80 %. Pour un classement en cinq stades, les résultats vacillent autour de 60-70 % selon les stades (Granfeldt & Nugues, 2007). Afin de mettre ces résultats en perspective, on peut comparer avec une expérience d'évaluation manuelle avec sept professeurs de FLE que nous avons effectuée (voir ci-dessous et Granfeldt & Ågren, sous presse). Dans cette étude le taux de convergence entre les professeurs était autour de 75 %, donc légèrement supérieur aux résultats obtenus pour la convergence homme-machine avec *Direkt Profil*.

#### **4. *Direkt Profil* – entre RAL et DDL**

Revenons à l'hypothèse avancée dans l'introduction, selon laquelle un logiciel comme *Direkt Profil* pourrait avoir un rôle médiateur à jouer entre les domaines des RAL et de la DDL. Nous aborderons surtout les possibles fonctions d'interface que *Direkt Profil* pourrait remplir dans le contexte de l'enseignement du FLE.

Il convient d'abord de souligner que les itinéraires d'acquisition ne doivent nullement être entendus comme la base du développement d'un curriculum en FLE ou même comme des repères pour une progression pédagogique dans l'enseignement. Les itinéraires d'acquisition sont le résultat d'un ajustement successif de la grammaire interne de l'apprenant dont l'état actuel, le stade de développement, est observable à travers certaines analyses et certains critères. Les RAL ont montré que les itinéraires d'acquisition sont communs aux apprenants (au moins ceux d'une certaine L1) et qu'ils sont indépendants de l'enseignement et au contexte d'appropriation. Pourtant, le rythme avec lequel les apprenants traversent les différentes étapes varie, entre autres, en fonction de la qualité et de la quantité d'input ainsi que de l'âge de l'apprenant. Il ne s'agit donc nullement de suggérer que l'enseignement suive explicitement l'ordre observé de l'acquisition des phénomènes grammaticaux.

L'intérêt principal des itinéraires d'acquisition et des stades de développement pour la DDL se trouve, à notre avis, plutôt dans l'activité évaluative, notamment l'évaluation des compétences linguistiques de l'élève dans un contexte didactique. On distingue traditionnellement trois types d'évaluations : 1) évaluation initiale (ou diagnostique), 2) évaluation formative, et 3) évaluation sommative. En acceptant l'autonomie et l'imperméabilité relative des itinéraires d'acquisition et des stades de développement ainsi que les rythmes individuels avec lesquels les apprenants les traversent, nous proposons que l'évaluation dite sommative se prête moins bien à l'appropriation des compétences linguistiques en langues étrangères. Dans l'état actuel des RAL, il nous paraît impossible de dire avec certitude qu'un apprenant se trouvera à tel ou tel niveau de compétence en langue étrangère après un certain nombre d'heures d'études. Les recherches très récentes sur l'interaction des facteurs externes (quantité et qualité d'input, *etc.*) et internes (âge, dispositions individuelles, motivation *etc.*) sur la route et le rythme d'acquisition de la langue étrangère permettront éventuellement d'éclaircir cette question très importante dans l'avenir (voir Ågren, Granfeldt & Thomas, sous presse pour une étude multifactorielle portant sur le français L2). C'est aussi la raison pour laquelle nous sommes de l'avis que les acquis des RAL sont plus intéressants pour l'évaluation diagnostique et formative, surtout que *Direkt Profil*, en principe, peut s'insérer dans diverses situations didactiques.

En DDL, les outils d'évaluation sont traditionnellement classés selon un axe qui va d'outils ouverts (pour l'écrit : rédaction de textes libre, résumés, synthèses, *etc.*) aux outils fermés (QCM, textes à trous). Les outils fermés sont typiquement utilisés pour évaluer des connaissances linguistiques précises aux niveaux grammatical et lexical alors que les outils ouverts, étant donné leur nature, sont souvent conçus pour une évaluation holistique de la compétence communicative.

L'une des caractéristiques de *Direkt Profil* est son degré d'ouverture et d'adaptation qui, grâce aux technologies de TAL, permet d'analyser et d'évaluer

systématiquement et objectivement certains aspects grammaticaux de tout texte rédigé par un apprenant de FLE. Ainsi, *Direkt Profil* est un exemple d'outil d'évaluation ouvert (appliqué sur la rédaction libre de l'élève), mais avec la capacité d'un type d'évaluation (systématicité, niveau de précision détaillé, automatique) qui est normalement associé aux outils fermés.

En outre, la validité de toute évaluation en contexte didactique est fondée sur la cohérence entre les activités proposées dans l'enseignement, le format et le type d'évaluation et la finalité du programme. Dans l'approche communicative qui domine actuellement la DDL, les activités pédagogiques et le contenu visent à améliorer la compétence communicative des élèves, c'est-à-dire leur capacité à utiliser la langue étrangère pour transmettre et comprendre des messages. L'intérêt de *Direkt Profil* est de permettre d'évaluer les performances réalisées dans la communication en temps réel, où les connaissances implicites de la langue étrangère deviennent visibles et possibles à modéliser.

Dans le but de tester l'intérêt concret des stades de développement et la performance de *Direkt Profil* pour l'évaluation en FLE, nous avons procédé à une recherche expérimentale (Granfeldt & Ågren 2008, sous presse). Nous avons demandé à sept professeurs expérimentés de FLE en Suède d'évaluer 40 textes du corpus CEFLE rédigés en français (dont 10 rédigés par des jeunes francophones) au niveau de la forme (grammaire, lexique, orthographe) et du contenu (aspects communicatifs). Les mêmes textes ont été analysés par *Direkt Profil* afin d'établir un taux de corrélation entre ces deux approches. Les résultats ont montré un haut degré de corrélation entre les évaluations faites par les professeurs et *Direkt Profil* (le coefficient de corrélation ( $r$ ) est supérieur à 0.8). Ce résultat vaut à la fois pour les évaluations portant sur la forme et sur le contenu. Sans entrer dans les détails ici (voir Granfeldt & Ågren, 2008, sous presse), notre interprétation des résultats est qu'il y a une convergence suffisante pour continuer ces recherches.

Nous insisterons, en conclusion, sur les aspects d'adaptation et d'ouverture de *Direkt Profil*. Le système est actuellement construit sur un corpus de récits d'apprenants suédophones. Dans cette configuration, au moins deux paramètres sont susceptibles d'influencer les résultats : le type de texte (récit) et la L1 des apprenants (le suédois). Grâce au module d'apprentissage automatique du logiciel, il est parfaitement possible d'entraîner les algorithmes sur des corpus où ces deux paramètres sont différents (p.ex. des textes descriptifs ou argumentatifs, ou avec d'autres langues sources) afin de produire, si besoin est, des classificateurs adaptés à d'autres contextes. Pour ne prendre qu'un exemple, les résultats de Granget (2006) discutés ci-dessus suggèrent par exemple que les apprenants allemands de FLE pourraient acquérir l'accord sujet-verbe plus rapidement que les apprenants suédophones. Le profil grammatical pour le stade 3 (voir Tableau 1) serait donc différent au moins sur ce point. C'est aussi dans la collecte de corpus didactiques

d'un haut niveau de spécificité (tâches et contextes précis, contraintes de production variées, etc.) que nous voyons encore la nécessité d'une future collaboration entre les RAL et la DDL.<sup>7</sup>

## Conclusion

Dans cet article, nous avons présenté *Direkt Profil*, un analyseur morphosyntaxique conçu pour analyser des textes de FLE, construit à l'aide de techniques de TAL établies et basées sur des RAL. Nous avons discuté l'hypothèse selon laquelle ce type de logiciel pourrait servir d'interface entre les RAL et la DDL, notamment dans les domaines de l'évaluation diagnostique et formative. Sur la base de textes rédigés librement par les élèves/apprenants, *Direkt Profil* est capable d'analyser un certain nombre de constructions grammaticales, correctes ou non, et ainsi montrer à l'enseignant et/ou à l'élève où il se situe dans son développement (son profil morphosyntaxique). Nous avons également suggéré qu'il existe plusieurs domaines où une collaboration étroite entre la DDL et les RAL seraient nécessaires pour répondre aux besoins futurs de ces projets scientifiques.

## Bibliographie

Ågren, M. (2008). À la recherche de la morphologie silencieuse : Sur le développement du pluriel en français L2 écrit. *Études Romanes de Lund*, 84. [Thèse de doctorat, Université de Lund]. Récupéré du site de l'université : <http://lup.lub.lu.se/luur/download?func=downloadFile&recordId=1267092&fileId=1267127>

Ågren, M., Granfeldt, J. & Thomas, A. (sous presse). Combined effects of age of onset and input on the development of different grammatical structures : A study of simultaneous and successive acquisition of French. *Linguistic Approaches to Bilingualism*.

Bartning, I. & Schlyter, S. (2004). Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies*, 14, 281-299.

Coste, D. (1992). Linguistique de l'acquisition et didactique des langues : repères pour des trajectoires. Dans R. Boucharde (Dir.) *Acquisition et enseignement/apprentissage des langues* (pp. 319-328), Grenoble : LIDILEM.

---

7. Suite à l'entrée successive des TIC dans la salle des classes, la mise en place de corpus didactiques (électroniques) est considérablement facilitée. Dans un projet en cours, nous avons collaboré avec un certain nombre d'écoles où les élèves rédigent des textes sur ordinateur par l'intermédiaire d'un site web développé par le groupe de chercheurs. Après la rédaction, les textes sont envoyés directement au serveur de l'institut de recherches pour stockage, annotation et analyses ultérieures. A notre avis, il faut envisager plus de projets de recherches collaboratifs entre enseignants, chercheurs en DDL et en RAL.

Devitt, S. (2004). Le cadre européen commun de référence pour les langues et les recherches sur l'acquisition des langues L1. Dans J.-C. Beacco, S. Bouquet et R. Porquier. *Niveau B2 pour le français. Textes et références* (pp. 9-16). Paris : Didier.

Fayol, M. (2003). L'acquisition/apprentissage de la morphologie du nombre. Bilan et perspectives. *Rééducation Orthographique*, 213, 151-166.

Fayol, M. & Jaffré, J.P. (2008). *Orthographier*. Paris : PUF.

Granfeldt, J. (2006). Evaluation du niveau lexical et grammatical à l'écrit en français langue étrangère : l'apport des analyses automatiques. *Revue française de linguistique appliquée*, 11, 103-119.

Granfeldt, J., Nugues, P., Person, E., Persson, L., Kostadinov, F., Ågren, M. & Schlyter, S. (2005). *Direkt Profil* : A system for evaluating texts of second language learners of French based on developmental sequences. *Proceedings of the second workshop on building educational applications using NLP* (pp. 53-60). Annual meeting of the association of Computational Linguistics, Ann Arbor, US.

Granfeldt, J. & Nugues, P. (2007). Evaluating stages of development in second language French : A machine learning approach. Dans J. Nivre, H.-J. Kaalep, K. Muischnek, & M. Koit (Dir.). *NODALIDA 2007 Conference Proceedings* (pp. 73-80). Tartu : Université de Tartu.

Granfeldt, J. & Ågren, M. (2008). Grammatisk profilanalys och lärarbedömning : En korrelationsstudie. Dans J. Granfeldt, G. Håkansson, M. Källkvist & S. Schlyter (Dir.) *Språkinläring, språkdidaktik och teknologi. Rapport från ASLAs höstsymposium 2007* (pp. 79-100). Lund : Association Suédoise de Linguistique Appliquée.

Granfeldt, J. & Ågren, M. (sous presse). SLA developmental stages and teachers' assessment of written L2 French : Exploring *Direkt Profil* as a diagnostic assessment tool. *Language Testing*, 31.

Granget, C. (2006). Développement de l'accord verbal avec un sujet pluriel dans les récits écrits d'apprenants germanophones scolarisés du français. Dans J. Granfeldt & S. Schlyter (Dir.) *Acquisition et production de la morphologie flexionnelle. Actes du Festival de la Morphologie*, PERLES no 20, Université de Lund.

Housen, A., Kemps, N. & Pierrard, M. (2007). Le développement de la morphologie verbale chez des apprenants avancés de FLE : apports et limites du contexte instructionnel, *Actes du Colloque International « Recherches en acquisition et en didactique des langues étrangères et secondes »*, Paris, Septembre 2006.

Housen, A., Kemps, N. & Pierrard, M. (2009). The use of verb morphology of advanced L2 learners and native speakers of French. Dans E. Labeau & F. Myles (Dir.). *The advanced learner varieties : The case of French* (pp. 41-61). Berlin : Peter Lang.

Huver, E. & Springer, C. (2009) *L'évaluation en langues*. Paris : Didier.

Trévisiol-Okamura, P. & Komur-Thillooy, G. (Dir.) (2011) *Discours, acquisition et didactique des langues. Les termes d'un dialogue*. Université de Haute-Alsace : Orizons.

Vergne, J. (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, 29 septembre 1999, Caen.

Véronique, D. (Dir.) (2000). Didactique des langues étrangères et recherches sur l'acquisition. *Études de Linguistique Appliquée*, 120.

Véronique, D. (Dir.) (2009). *L'acquisition de la grammaire du français, langue étrangère*. Paris : Didier.