



**HAL**  
open science

## Diagnostic et désambiguïisation des erreurs de flexion dans un outil ELIAO

Ivan Šmilauer

► **To cite this version:**

Ivan Šmilauer. Diagnostic et désambiguïisation des erreurs de flexion dans un outil ELIAO. Colloque international des jeunes chercheurs en Didactique des Langues et en Linguistique (CEDIL 2010), LIDILEM, Jun 2010, Grenoble, France. pp.19-26. hal-01375626

**HAL Id: hal-01375626**

**<https://inalco.hal.science/hal-01375626v1>**

Submitted on 20 Jan 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## DIAGNOSTIC ET DESAMBIGUÏSATION DES ERREURS DE FLEXION DANS UN OUTIL ELIAO

Ivan ŠMILAUER

[smilauer@cetlef.fr](mailto:smilauer@cetlef.fr)

INALCO, LaLIC-CERTAL

49bis, avenue de la Belle Gabrielle

75012 Paris

France

Abstract : The analysis of learners' language and the generation of a linguistically motivated feedback is one of the main applications of NLP in Intelligent Computer Assisted Language Learning (ICALL). This paper presents a diagnosis of inflectional errors in Czech implemented on [www.CETLEF.fr](http://www.CETLEF.fr) and its effectiveness on a corpus of authentic productions.

### 1. Objectifs et limites actuelles du diagnostic automatique des erreurs en L2

Les outils d'enseignement des langues "intelligemment" assisté par ordinateur (ELIAO)<sup>1</sup> tentent de mettre au profit des apprenants les techniques du traitement automatique des langues (TAL). Depuis une vingtaine d'années, le diagnostic des erreurs dans les productions langagières et la génération automatique d'une rétroaction didactique sont un des principaux objectifs du domaine, voir notamment Nerbonne (2003), Heift & Schulze (2007), Meurers (2009). Le défi n'est pas seulement d'identifier et de corriger les erreurs (d'orthographe, de grammaire, de sens, de style, etc.), mais également de les comprendre – établir un diagnostic et fournir un retour à l'apprenant. La complexité de cette tâche requiert la modification des outils existants, l'adoption de nouvelles stratégies et le développement de nouveaux outils.

L'emploi d'un dispositif de correction générique sur des textes en L2 amène au moins deux problèmes. Il s'agit d'abord du traitement des **erreurs morphologiques** (par ex. *writed*, *carnavaux*, *chevals* etc.), rares chez les locuteurs natifs mais typiques (et très fréquentes dans les langues flexionnelles) pour les apprenants de L2. Ces formes, inexistantes dans la langue cible, sont interprétées par des correcteurs d'orthographe classiques comme des erreurs de frappe ou d'orthographe, voir par ex. Rimrott & Heift (2008). Étant donné qu'il s'agit des formes avec une évidente motivation linguistique (surgénéralisation, non-respect des alternances et des idiosyncrasies, erreurs dans l'attribution des types de flexion, etc.), une information importante et utile pour le diagnostic reste ainsi inexploree.

En ce qui concerne la correction de grammaire, il faut admettre que les méthodes existantes de l'**analyse syntaxique** sont imparfaites même avec des entrées standards, ce qui s'amplifie nécessairement dans le cas des textes contenant des erreurs. Les obstacles les plus importants sont l'ambiguïté des formes lexicales, l'ordre libre des éléments de phrase dans certaines

---

<sup>1</sup> Dans le milieu anglophone, les outils et la recherche connexe sont appelés Intelligent Computer Assisted Language Learning (ICALL), dans le milieu francophone, ce sont les termes ELAO ou ALAO (Enseignement ou Apprentissage des Langues Assistés par Ordinateur) intégrant le TAL qui sont employés le plus souvent, voir par ex. Hamel & Vandeventer (1998), Antoniadis et al. (2005).

langues, l'attachement des syntagmes prépositionnels, etc. Les analyseurs existants peuvent être modifiés afin de produire un diagnostic des constructions agrammaticales, cependant les méthodes proposées sont loin d'être suffisamment fiables et les résultats sont assez bruités, voir par ex. L'Haire & Faltin (2003) ou Schneider & McCoy (1998). Un problème supplémentaire spécifique à l'analyse des textes en L2 est la présence des erreurs dues à l'homophonie (par exemple *ont y va!*), voir Faltin (2003).

Tandis que des imperfections peuvent être tolérées dans des applications destinées aux locuteurs natifs, elles sont difficilement acceptables dans des outils avec une visée didactique. Actuellement, la stratégie la plus convenable semble être de se limiter à des techniques bien maîtrisées (correction orthographique et analyse morphologique) et de les employer d'une manière ciblée pour un diagnostic fiable, voir par ex. Kraif et al. (2004) ou Blanchard (2007). Bien que le défi d'ELIAO soit le traitement de la complexité linguistique des productions libres, il est donc plus raisonnable de se concentrer sur l'analyse des formes hors contexte ou de traiter uniquement des productions issues de tâches fermées comme des exercices de type phrase à trous (nombre limité de productions correctes), voir par ex. Kraif & Ponton (2007).

## 2. Diagnostic des erreurs de flexion dans les tâches fermées

En accord avec l'approche mentionnée ci-dessus, nous avons développé un diagnostic des erreurs de flexion issues des tâches fermées, ciblées sur la flexion nominale du tchèque, voir Šmilauer (2008). Ce diagnostic est implémenté sur la plateforme [www.CETLEF.fr](http://www.CETLEF.fr), qui propose des exercices de déclinaison avec des phrases à trous et un feedback automatique sur les productions des apprenants<sup>2</sup>.

### 2.1. Aperçu de la déclinaison tchèque

Le tchèque possède une flexion nominale très riche : sept cas, deux nombres, quatre genres, 14 types paradigmatiques avec de nombreux sous-types et des exceptions et de nombreuses alternances vocaliques et consonantiques, qui se produisent en fonction des propriétés morpho-phonologiques du radical et de la désinence casuelle, par ex. *dům* > *dom-u*, *kočka* > *kočc-e*, etc.

Le tableau 1 donne un aperçu de cette richesse et il permet de constater une forte homonymie dans le système des désinences casuelles (une forme a plusieurs fonctions) qui est une propriétés importante à prendre en compte pour la conception du diagnostic.

	singulier			
	masculin animé	masculin inanimé	féminin	neutre
<b>nominatif</b>	-Ø, -a, -e, -ě	-Ø	-Ø, -a, -e, -ě	-e, -ě, -í, -o
<b>génitif</b>	-a, -e, -ě, -y	-a, -e, -ě, -u	-e, -ě, -i, -y	-a, -e, -ě, -í
<b>datif</b>	-i, -ovi, -u	-i, -u	-e, -ě, -i	-i, -í, -u
<b>accusatif</b>	-a, -e, -ě, -u	-Ø	-Ø, -i, -u	-e, -ě, -í, -o
<b>vocatif</b>	-e, -i, -o, -u	-e, -ě, -i, -u	-e, -ě, -i, -o	-e, -ě, -í, -o
<b>locatif</b>	-i, -ovi, -u	-e, -ě, -i, -ovi, -u	-e, -ě, -i	-e, -ě, -i, -í, u
<b>instrumental</b>	-em, -ěm, -ou	-em, -ěm	-í, -ou	-em, -ěm, -ím

Tableau 1 - Désinences casuelles du singulier des substantifs et leurs fonctions

<sup>2</sup> CETLEF (Connaître, Comprendre et Corriger les Erreurs en Tchèque Langue Étrangère pour les Francophones) est une application web dynamique (PHP, MySQL, XHTML) disponible librement sur <http://www.cetlef.fr> à partir de juin 2008.

Pour les besoins du diagnostic, un modèle formel de la déclinaison tchèque a été élaboré. Ce modèle est composé des listes de désinences définissant les types paradigmatiques (125 au total) et des règles contextuelles pour la réalisation des alternances (32 règles d'alternance).

Contrairement à la morphologie, le système orthographique du tchèque est presque phonologique et il manifeste un très bas niveau d'homophonie ou d'homographie.

## 2.2. Tâche de déclinaison

La tâche de l'apprenant est de décliner un substantif, un adjectif ou un pronom en fonction de son contexte syntaxique dans une phrase à trous (voir la figure 1). Lors de la création des exercices, chaque **forme requise (FR)** est annotée par un analyseur morphologique qui détermine sa catégorie lexicale, les valeurs des catégories grammaticales (*cas*, *num*, *gen*), son type paradigmatique (*pdgm*) et la réalisation d'une éventuelle alternance morphématique vocalique ou consonantique (*alt*).

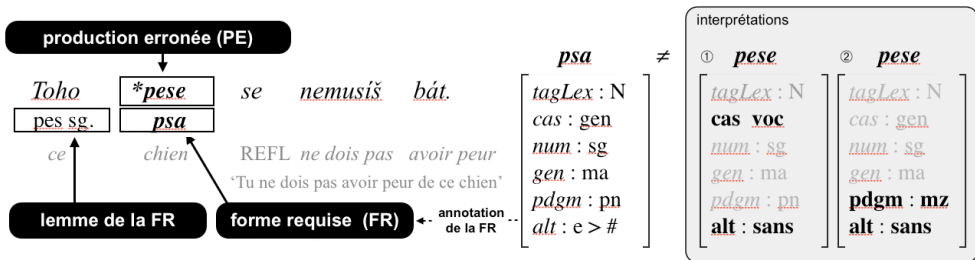


Figure 1 - Une tâche de déclinaison, l'annotation et les interprétations d'une PE

Si la production saisie par l'apprenant n'est pas égale à la FR, elle est considérée comme une **production erronée (PE)** et le diagnostic est enclenché pour déterminer sa nature.

## 2.3. Description du diagnostic

En s'inspirant des travaux sur l'acquisition de L2 et le traitement de la morphologie du point de vue psycho-computationnel, voir par ex. Pirelli (2007), nous partons de l'hypothèse qu'une PE dans une tâche de déclinaison n'est pas une forme aléatoire mais calculable et elle peut être reproduite artificiellement. Une approche similaire a été proposée dans un prototype de correcteur basé sur la prévisibilité des erreurs dans la flexion verbale de l'allemand par Rimrott (2003) ou de l'espagnol par Mansilla (2004).

Nous supposons donc qu'une PE peut être générée à partir du lemme de la FR à l'aide des trois opérations suivantes :

1. l'ajout des désinences casuelles inappropriées au radical de la forme requise
2. le non-respect des alternances morphématiques
3. la violation des règles de graphie (diacritique)<sup>3</sup>

Le diagnostic d'une PE est effectué en deux étapes : la **recherche des interprétations** possibles et le **filtrage de ces interprétations**.

### 2.3.1. Recherche des interprétations

La procédure tente d'abord d'éliminer des formes qui ne devraient pas avoir une motivation morphologique (notamment des formes saisies par les apprenants uniquement pour avancer dans les exercices et n'ayant rien de commun avec la FR). Ces **tests non morphologiques**

<sup>3</sup> Nous incluons la gestion des erreurs de diacritique dans le traitement morphologique car l'absence d'un signe diacritique peut marquer une fonction différente (par ex. *růži* et *růží* sont deux formes casuelles du substantif *růže*).

vérifient la présence des caractères étrangers, la différence dans longueur de la FR et de la PE et la distance maximale de Levenshtein<sup>4</sup> admissible entre deux formes fléchies d'un substantif (égale à 4 en tchèque, par ex. pour *pes* > *psovi*).

Si aucun de ces tests ne trouve une interprétation, le diagnostic se poursuit par la recherche des interprétations morphologiques d'une PE. Celle-ci est effectuée par la comparaison de la PE avec des **formes hypothétiques (FH)**, générées à partir du radical de la FR à l'aide du modèle formel de la déclinaison tchèque décrit ci-dessus. Si une correspondance est trouvée, la différence entre la FH et la FR est retenue comme une possible **interprétation morphologique** de la PE (voir par ex. les interprétations de la forme *pese* sur la figure 1). La taxonomie des **erreurs morphologiques** est basée sur les valeurs des attributs qui diffèrent pour la FR et la FH. (par ex. *erreur de cas et d'alternance* et *erreur de type paradigmatique et d'alternance* sur la figure 1).

Afin de limiter le nombre des interprétations possibles en éliminant celles qui seraient improbables du point de vue de l'activité langagière de l'apprenant (par exemple une *erreur de cas, de nombre et de genre*), il existe dans la version actuelle du diagnostic une **restriction sur la génération des FH**. Celles-ci ne sont générées que par les désinences au sein du type paradigmatique de la FR. En dehors de ce type (les erreurs de type paradigmatique et de genre), les valeurs de cas et de nombre restent inchangées par rapport à la FR et les désinences sont choisies uniquement à partir des paradigmes ayant les mêmes propriétés morpho-phonologiques que le paradigme de la FR (les nominatifs des deux types doivent être terminés par une désinence identique). Ceci permet de rendre compte des erreurs où l'apprenant attribue un type paradigmatique erroné à un mot à cause de sa ressemblance formelle avec un substantif d'un autre type (par ex. le substantif neutre *moře* 'mer' est décliné par erreur d'après un type à cause de la désinence homonyme *-e*, commune pour le nominatif des types féminins et neutres). Comme nous le verrons plus bas, cette restriction permet de réduire drastiquement le nombre des interprétations d'une PE, mais elle limite la puissance du diagnostic dans quelques cas extrêmes.

### 2.3.2. Filtrage des interprétations

À cause de l'homonymie des désinences casuelles, il y a souvent plusieurs interprétations de la PE (voir aussi la figure 3). Ces interprétations, trouvées lors de la génération de FH, sont groupées d'après le type d'erreur (par ex. toutes les interprétations avec *une erreur de cas* sont considérées désormais comme une seule interprétation) pour entrer en tant qu'une unité dans la procédure de filtrage qui a pour l'objectif de choisir les interprétations les plus plausibles. Deux critères, fondés sur des hypothèses posées lors du développement, sont employés pour le filtrage :

- **Nombre d'attributs atteints par l'erreur** : l'interprétation qui contient le moins d'attributs atteints par l'erreur est retenue comme la plus plausible (par ex. une interprétation *erreur de cas* est considérée comme plus plausible qu'une interprétation *erreur de cas et de nombre*).
- **Pertinence du trait atteint** : certains traits sont considérés comme plus pertinents que d'autres : l'erreur de nombre précède l'erreur de genre et l'erreur de cas et une interprétation avec une erreur de diacritique est considérée comme moins pertinente qu'une interprétation identique sans cette erreur.

Le diagnostic final retient au maximum deux interprétations car souvent il n'est pas possible de décider laquelle est la meilleure. Les interprétations retenues sont inscrites dans la base de

<sup>4</sup> La distance de Levenshtein est utilisée pour la mesure de la proximité entre deux chaînes. Il s'agit d'une valeur numérique exprimant le nombre minimal d'opérations sur des caractères d'une chaîne (ajout, suppression, substitution) nécessaires pour la rendre identique à une autre chaîne.

données pour spécifier les PE et elles sont utilisées pour le formatage du feedback destiné à l'apprenant. Le feedback a la forme d'un message informant l'apprenant sur la nature de l'erreur. Pour la PE dans la figure 1, le feedback a la forme suivante : « *Il y a deux interprétations possibles de votre erreur : (1) Le cas et le nombre sont corrects mais peut être que vous confondez le type de déclinaison et que vous déclinez le mot d'après le type muž. (2) Le nombre et le genre sont corrects, mais il y a une erreur de cas. Concernant votre forme, vous n'avez pas effectué l'alternance e > #.* »

### 3. Évaluation du diagnostic sur un corpus de productions

Le diagnostic a été testé sur 11 734 productions issues de 525 tâches différentes et saisies par 213 apprenants francophones<sup>5</sup>. 3 859 productions (32,9 %) sont des PE.

#### 3.1. Propriétés formelles des PE

La répartition des valeurs de la distance de Levenshtein pour toutes les PE est représentée sur la figure 2 (la courbe pointillée représente la mesure prise sur les formes sans diacritique).

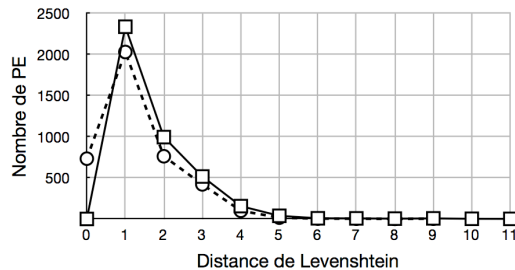


Figure 2 - Distance de Levenshtein entre les FR et les PE recueillies

Plus de 50 % des PE diffèrent uniquement par un seul caractère et la quasi-totalité des PE a une valeur inférieure à 4, qui est la distance maximale entre deux formes nominales fléchies en tchèque.

La fig. 3 illustre l'impact des restrictions introduites dans la génération de FH. Elle montre le nombre de différentes interprétations (avant le filtrage) pour toutes les PE dans le corpus, obtenu par deux variantes du diagnostic.

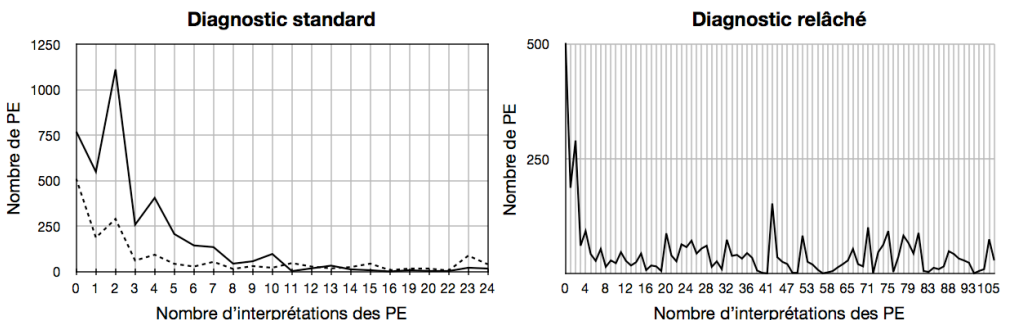


Figure 3 - Nombre des interprétations (FH = PE) générées par deux types de diagnostic

<sup>5</sup> Le nombre de productions par tâche est variable car tous les apprenants n'ont pas effectué toutes les tâches qui sont présentées en un parcours fixe, construit d'après la progression didactique traditionnelle reflétant la difficulté des différents phénomènes.

Le diagnostic standard ne génère que les formes hypothétiques dans le cadre de la restriction formelle posée sur la génération, voir plus haut, section 2.3.1. Le nombre maximal d'interprétation par PE est 24 et la moyenne est de 3,1.

Le diagnostic relâché (représenté également par une courbe pointillée sur le graphique de gauche) ne respecte pas la restriction sur la génération, ce qui multiplie le nombre des FH correspondantes par 10 (35,8 interprétations par PE en moyenne) avec le nombre maximal d'interprétations de 105 (!). Le seul avantage du diagnostic relâché est la baisse du nombre des PE sans interprétation d'un tiers environ.

### 3.2. Fréquence des types d'erreurs

2 978 productions (77,2 % de toutes les PE) ont été interprétées comme une erreur morphologique (voir tableau 2).

une interprétation retenue		deux interprétations retenues		
type d'erreur	%	type d'erreur		%
locale	28,9	cas	genre	7,0
cas	26,8	locale	cas	5,4
nombre	9,8	type pdgm	genre	2,7
genre	5,4	nombre	genre	2,1
type pdgm	4,3	nombre	type pdgm	1,2
cas et nombre	3,7	locale	nombre	0,9
<b>Total</b>	<b>78,9</b>	cas	type pdgm	0,5
		cas et nombre	genre	0,4
		cas et nombre	type pdgm	0,3
		locale	cas et nombre	0,3
		locale	cas et genre	0,3
		<b>Total</b>		<b>21,1</b>

erreurs de diacritique et d'alternance	
type d'erreur	%
diacritique	32,6
alternance	12,1

Tableau 2 - Fréquence des différents types d'erreurs

Pour la plupart des erreurs morphologiques (78,9 %), un diagnostic avec une seule interprétation a été établi. L'erreur la plus fréquente est l'erreur dite *locale* : les valeurs des catégories de cas, de nombre et de genre ne sont pas atteintes et la PE contient une erreur de diacritique ou d'alternance<sup>6</sup>. L'erreur de cas est le deuxième type le plus fréquent. Les PE avec un diagnostic à deux interprétations sont moins nombreuses, l'ambiguïté la plus fréquente réside premièrement entre une *erreur de cas* et de *genre* et deuxièmement, entre une *erreur locale* (le plus souvent d'une erreur de diacritique) et une *erreur de cas*.

Parmi les 22,9 % des PE restantes, 4,4 % ont été identifiées comme une erreur sans motivation morphologique et pour 18,5 % le diagnostic n'a trouvé aucune interprétation.

## 4. Propositions d'amélioration du diagnostic

Il est possible d'améliorer la performance du diagnostic de deux manières. D'une part, nous pouvons modifier les restrictions sur la génération des formes hypothétiques afin d'obtenir moins de PE sans interprétation; d'autre part, il serait possible d'améliorer le filtrage des interprétations ambiguës en ajoutant des critères supplémentaires permettant de choisir l'interprétation la plus plausible.

<sup>6</sup> Les fréquences de ces types d'erreurs ont été mises à part pour des raisons de place et de clarté, car elles devraient accompagner tous les types d'erreur, ce qui multiplierait la taille du tableau par 3.

### 4.1. *Analyse des erreurs sans interprétation*

Les PE sans interprétation représentent un cinquième de toutes les PE. Une analyse de ses productions a révélé trois groupes de formes relativement homogènes.

La partie majoritaire de ces PE sont des **fautes de frappe** classiques (ajout, effacement, substitution d'un ou plusieurs caractères) avec la distance de Levenshtein souvent égale à 1. Ces formes ne semblent avoir aucune motivation linguistique et dans la plupart des cas, elles sont liées au médium informatique (clavier d'un ordinateur). Il serait possible de remplacer ces formes par la FH la plus proche en terme de distance de Levenshtein et de poursuivre normalement le diagnostic avec cette nouvelle PE.

Le deuxième groupe est celui des **erreurs morphologiques non identifiées** à cause des restrictions posées sur la génération des FH. L'hypothèse sur la restriction au cas et au nombre de la FR semble être justifié, néanmoins la restriction sur les propriétés morpho-phonologiques des paradigmes à partir desquels les désinences peuvent être empruntées pour la génération des FH semble être trop forte. En effet, le diagnostic ne reconnaît pas les formes erronées à cause du choix d'une désinence typique pour un type paradigmatic sans qu'il y ait une ressemblance formelle entre le mot décliné et le type choisi (par ex. la PE *městovi* au lieu de *městu*, est créée à partir du neutre *město* 'ville' par une désinence typique pour le datif singulier des masculins animés du type *pán*). Un relâchement partiel des contraintes sur la génération des FH devrait permettre d'interpréter ces formes.

Le dernier groupe pourrait être nommé comme celui des **erreurs « lexicalement suspectes »**. Il s'agit des formes qui sont morphologiquement et lexicalement motivées par le lemme de la FR mais qui ressemblent en même temps à un autre lexème (il s'agit par ex. des diminutifs : *koláček* 'petit gâteau' au lieu de *koláči* ; des adjectif dérivés : *francouzský* 'français' au lieu de *Francouz* 'un Français' ; des féminins : *profesorka* au lieu de *profesor* etc.) Il semble qu'il est impossible d'interpréter ces formes avec la méthode de génération des FH actuelle, basée sur des critères purement flexionnels.

### 4.2. *Enrichissement du filtrage des interprétations*

La seconde manière consiste dans le raffinement des critères qui pourraient être utilisés pour le choix de la meilleure interprétation pour des PE ambiguës. Bien que l'adéquation du critère du nombre minimal de traits atteints par l'erreur semble être vérifiée sur les PE observées, le critère de la pertinence des traits est plus discutable et l'étude des diagnostics avec deux interprétations concurrentes montre qu'il n'est pas possible (à part la discrimination des interprétations par la présence d'une erreur de diacritique) d'établir une échelle de pertinence applicable dans toutes circonstances.

Étant donné que les formes fléchies sont des expressions des fonctions syntaxiques, l'introduction des **informations concernant la valence** de l'élément régisseur de la FR pourrait jouer un rôle dans l'attribution d'une préférence à l'interprétation *erreur de cas* d'une PE ambiguë. Dans le cas du tchèque, il serait possible d'enrichir la description des tâches par les informations issues du lexique électronique de valences verbales VALLEX 2.0, voir Žabokrtský & Lopatková (2007). Cependant, les cas d'ambiguïté où cette information pourrait jouer un rôle déterminant (par exemple la confusion des cadres verbaux de deux verbes homonymes) sont très rares dans le corpus. L'effort investit dans cet enrichissement pourrait donc être effectif uniquement si l'information d'ordre syntaxique enrichissait également le feedback destiné à l'apprenant : au lieu de signaler uniquement une erreur de cas, il serait possible d'attirer son attention sur le contexte syntaxique de la FR.

Une autre aide à la désambiguïsation serait l'introduction de l'information sur le **genre de l'équivalent français de la FR**. Les erreurs de genre dues à l'interférence de la langue maternelle sont assez fréquentes et dans le cas d'une PE ambiguë, cette interprétation pourrait



être privilégiée au détriment de l'autre. L'étude des PE montre que la quasi-totalité des interprétations *erreur de genre* peut être admise comme la plus plausible pour les substantifs avec le genre différent dans les deux langues. Cette différence serait également une information utile pour le feedback destiné à l'apprenant.

## 5. Conclusion

Avec environ 80 % de réussite, le diagnostic des erreurs de flexion construit uniquement sur la caractérisation morphologique de l'écart entre la production erronée et la forme requise offre une solution fonctionnelle et relativement fiable. Le corpus d'erreurs recueillies a permis d'identifier les points faibles actuels et d'envisager des améliorations dans le cadre donné. De plus, la connaissance des différents types d'erreurs, de leur fréquence et des problèmes liés à leur ambiguïté pourrait être utile pour la modification des outils d'analyse plus complexes. À condition que nous disposions de la forme correcte et de ses catégories grammaticales, la méthode présentée ci-dessus pourrait être employée dans l'analyse des productions libres.

## Références

- ANTONIADIS, Georges et al. (2005). Modélisation de l'intégration de ressources TAL pour l'apprentissage des langues : la plateforme MIRTO. *ALSIC : Apprentissage des Langues et Systèmes d'Information et de Communication*, 8:65–79.
- BLANCHARD, Alexia (2007). Analyse morphologique des réponses d'apprenants, in *Actes de TALN-RECITAL 2007, Toulouse 5-8 juin*:437–446.
- HAMEL, M.-J., and VANDEVENTER, A. (1998). Fipsgram : un analyseur syntaxique dans un contexte d'ELAO, in *Le traitement automatique du langage et les applications industrielles TAL + AI 98, I*, 18–24.
- HEIFT, Trude & SCHULZE, Mathias (2007). *Errors and Intelligence in Computer-Assisted Language Learning: Parsers and Pedagogues*. UK : D. Reidel Publishing Company.
- KRAIF, Olivier et al. (2004). NLP tools for CALL: the simpler, the better, in *Proceedings of InSTIL / ICALL2004 – NLP and Speech Technologies in Advanced Language Learning Systems (Venice, 17-19 June)*:37-40.
- KRAIF, Olivier & PONTON, Claude (2007). Du bruit, du silence et des ambiguïtés : que faire du TAL pour l'apprentissage des langues ? in *Actes de TALN 2007, Toulouse, France*.
- L'HAIRE, Sébastien & VANDEVENTER-FALTIN, Anne (2003). Error diagnosis in the Freetext project. *CALICO* 20, 3:481–495.
- MANSILLA, J. R. Z. (2004). Text generators, error analysis and feedback, in *Proceedings of InSTIL/ICALL2004 – NLP and Speech Technologies in Advanced Language Learning Systems*. Venice.
- MEURERS, Detmar (2009). On the Automatic Analysis of Learner Language. Introduction to the Special Issue. *CALICO Journal*, 26(3):469–473.
- NERBONNE, John (2003). Natural language processing in computer-assisted language learning, in *The Oxford Handbook of Computational Linguistics*, Mitkov, Ruslan (Ed). Oxford University Press, 670–698.
- PIRELLI, Vito. (2007). Psycho-computational issues in morphology learning and processing: An ouverture. *Lingue e Linguaggio*, VI(2):131–138.
- RIMROTT, Anne (2003). A spell checking algorithm for treating predictable verb inflection mistakes made by non-native writers of German, in *Term Paper for LING 807 – Computational Linguistics at Simon Fraser University*. Burnaby, Canada.
- RIMROTT, Anne, and HEIFT, Trude (2008). Evaluating automatic detection of misspellings in german. *Language Learning & Technology* 12, 3:73–92.

- 
- ŠMILAUER, Ivan (2008). *Acquisition du tchèque par les francophones : analyse automatique des erreurs de déclinaison*. Thèse PhD, INALCO / Faculté des lettres de l'Université Charles, Paris / Prague.
- ŽABOKRTSKÝ, Zdeněk & LOPATKOVÁ, Markéta (2007). Valency Information in Vallex 2.0. Logical Structure of the Lexicon. *Prague Bulletin of Mathematical Linguistics*, 87:41–60.