Adaptation par enrichissement terminologique en traduction automatique statistique fondée sur la génération et le filtrage de bi-segments virtuels

Christophe Servan, Marc Dymetman

Adaptation par enrichissement terminologique en traduction automatique statistique fondée sur la génération et le filtrage de bi-segments virtuels

Christophe Servan[1, 2] Marc Dymetman[2]

(1) LIG équipe GETALP , 41 rue des mathématiques, BP 53 38041 Grenoble Cedex 9
(2) Xerox Research Centre Europe, 6 chemin de Maupertuis, 38240 Meylan

christophe.servan@imag.fr, marc.dymetman@xrce.xerox.com

Résumé. Nous proposons des travaux préliminaires sur une approche permettant d'ajouter des termes bilingues à un système de Traduction Automatique Statistique (TAS) à base de segments. Ces termes sont, non seulement, inclus individuellement, mais aussi avec des contextes induits autour de ces mots. Tout d'abord nous générons ces contextes en généralisant des motifs (ou patrons) observés pour des mots de même nature syntaxique dans un corpus bilingue. Enfin, nous filtrons les contextes qui n'atteignent pas un certain seuil de confiance, à l'aide d'une méthode de sélection de bi-segments inspirée d'une approche de sélection de données, précédemment appliquée à des textes bilingues alignés.

Abstract. We propose a technique for adding bilingual terms to a phrase-based SMT system which includes not only individual words, but also induces phrasal contexts around these words. We first generate these contexts by generalizing patterns observed for similar words in a bilingual corpus, but then filter out those contexts that fall below a certain confidence threshold, based on an original phrase-pair selection process inspired by existing sentence selection techniques.

Mots-clés : Traduction Automatique Statistique, Génération Automatique de Texte, contexte phrastique, terminologie bilingue.

Keywords: Statistical Machine Translation, Natural Language Generation, phrasal context, bilingual terminology.

1 Introduction

La plupart des recherches concernant l'*adaptation* au domaine en Traduction Automatique Statistique (TAS) se basent sur un entrainement à partir d'un corpus bilingue de ce domaine (le plus souvent conjointement avec un corpus hors-domaine plus important, voir section 5) mais n'utilisent pas directement de terminologie spécifique à ce domaine. Par exemple, la traduction en français du mot anglais « layer » sera « calque » dans le domaine graphique, mais « couche » dans le domaine mathématique ou encore « niveau » si on est dans un contexte administratif.

Une solution est d'ajouter un dictionnaire bilingue spécifique à la table de bi-segments du modèle de traduction. Mais le fait d'ajouter des mots individuels sans aucun contexte (c'est-à-dire des mots individuels par opposition à des segments contenant ces mots) ne représente pas une utilisation optimale des capacités d'un système de TAS à base de segments [TAS-BS = PB-SMT en anglais] : un système limité à des bi-segments unigramme/unigramme est typiquement inférieur à un système utilisant des bi-segments plus étendus.

Ce type d'approche fondée sur l'enrichissement contextuel liée à une terminologie n'a que peu d'influence sur les résultats, de manière générale. Cependant, dans l'optique de fournir une traduction à post-éditer à un traducteur, ce type d'enrichissement est précieux et permet de gagner du temps de post-édition. Cette approche est donc destinée à un système de traduction automatique dans le cadre d'un processus de traduction assistée par ordinateur.

La proposition principale de cet article consiste à ajouter des entités nommées provenant d'un dictionnaire bilingue à la table de bi-segments, non directement, mais en reconstituant leurs contextes potentiels. Ces contextes potentiels sont obtenus par copie de contextes entourant des entités nommées similaires représentés dans la table originelle de bi-segments. Les bi-segments "virtuels" ainsi produits peuvent être sur-générés, c'est pourquoi nous proposons de leur appliquer certaines techniques de filtrage. Les travaux préliminaires présentés dans cet article proposent d'évaluer l'approche proposée

sur une terminologie associées aux entités nommées, dans l'optique de l'étendre dans de futurs travaux à tous types de mots (noms, adjectifs…)

2 Approche

Notre approche pour générer les contextes peut se décrire à l'aide de la « règle de déduction » suivante :

$$\frac{\alpha\beta\gamma \leftrightarrow_{pt} \alpha'\beta'\gamma'; \delta \leftrightarrow_{lex} \delta'; \beta, \delta : T; \beta', \delta' : T'}{\alpha\delta\gamma \leftrightarrow_{tbg} \alpha'\delta'\gamma'} \quad (1)$$

Dans cette règle, chaque lettre grecque dénote une suite de mots, et T, T' sont des types ; pt est la table de bi-segments (standard) originelle extraite du corpus bilingue. lex est une table de traduction fournie par l'utilisateur qui définit les correspondances terminologiques du domaine. Quant à tbg, c'est la table de bi-segments « généralisés » obtenue par application des déductions.

La règle dit que si $\alpha\beta\gamma \leftrightarrow_{pt} \alpha'\beta'\gamma'$ est une entrée dans la table originelle et si $\delta \leftrightarrow_{lex} \delta'$ est une correspondance terminologique, où β et δ sont du même type T (resp. β', δ' et T'), alors nous pouvons générer une nouvelle entrée $\alpha\delta\gamma \leftrightarrow \alpha'\delta'\gamma'$ où δ remplace β et respectivement δ' remplace β'. L'entrée nouvellement générée à l'aide de cette règle de déduction est alors ajoutée à la table de bi-segments généralisée ($\alpha\delta\gamma \leftrightarrow_{tbg} \alpha'\delta'\gamma'$).

2.1 Exemple : Entités Nommées

Dans ces travaux préliminaires, nous illustrons notre approche par un exemple de type terminologique : les noms de pays (c.-à-d. une classe d'entités nommées). Supposons que certains pays soient rarement ou pas du tout mentionnés dans notre corpus d'entrainement anglais-français, mais que nous ayons un dictionnaire qui nous donne leurs traductions lexicales. Par exemple, « Ecuador » apparait environ seulement 100 fois dans le corpus Europarl (Koehn, 2005), alors que « Germany » apparait environ 60 fois plus souvent. Nous pouvons considérer que les contextes linguistiques observés autour des pays peu représentés sont trop peu nombreux pour être fiables, et notre méthode consiste à tenter de transposer les contextes concernant les pays bien représentés aux pays peu représentés.

La première étape de notre approche consiste à identifier les noms de pays dans le corpus d'apprentissage. Une fois que les entités nommées de type pays sont identifiées, nous les remplaçons par un « marqueur » comme indiqué dans la Table 1. Pour plus d'efficacité, le processus d'extraction de patrons est effectué sur les bi-segments déjà extraits du corpus d'entrainement. Avec notre exemple « Ecuador », nous générons un nouveau bi-segment en remplaçant dans le patron *@COUNTRY@ is ||| L' @COUNTRY@ est* le marqueur source avec le terme « Ecuador » et le marqueur cible par sa traduction « Équateur ».

England is	|||	L' Angleterre est
Spain is	|||	L' Espagne est
Italy is	|||	L' Italie est
@COUNTRY@ is	|||	L' @COUNTRY@ est

@COUNTRY@ is	|||	L' @COUNTRY@ est
Ecuador is	|||	L' Équateur est

TABLE 1 – Exemple d'extraction du patron *@COUNTRY@ is ||| L' @COUNTRY@ est* (tableau de gauche) et de son application pour un même couple terminologique « Ecuador : Équateur ». (tableau de droite)

Le processus de génération de nouvelles entrées ($\alpha\delta\gamma \leftrightarrow \alpha'\delta'\gamma'$) peut générer des erreurs, au cas où les contextes virtuels générés ne sont pas compatibles avec les termes considérés.

Ainsi, plusieurs des segments virtuels générés pour le terme « Ecuador », illustrés dans le tableau 2 sont erronés : en français, l'article « Le » doit être élidé en « L' » devant une voyelle. Un grand nombre de problèmes de ce type peuvent apparaitre et c'est pourquoi un processus de filtrage doit être appliqué. Ce processus de filtrage est une contribution centrale de cet article décrit dans la section 3.

@COUNTRY@ is	|||	L' @COUNTRY@ est	from @COUNTRY@ ,	|||	des @COUNTRY@ ,
Ecuador is	|||	L' **Équateur** est	from **Ecuador** is	|||	des **Équateur** ,
@COUNTRY@ is	|||	Les @COUNTRY@ sont	from @COUNTRY@ ,	|||	de la @COUNTRY@ ,
Ecuador is	|||	Les **Équateur** sont	from **Ecuador** ,	|||	de la **Équateur** ,

TABLE 2 – Exemples de bi-segments générés qui peuvent contenir des erreurs (i.e. « from Ecuador » → « de la Équateur »).

3 Le processus de filtrage

Pour réaliser le filtrage, nous proposons d'utiliser une technique basée sur la différence de scores d'entropie croisée, inspirée par des approches récentes en sélection de données pour l'adaptation au domaine en TAS. Ces techniques de sélection sont appliquées soit au corpus cible uniquement (Moore & Lewis, 2010) (filtrage monolingue), soit conjointement aux corpus source et cible (Axelrod *et al.*, 2011) (filtrage bilingue).

Filtrage monolingue Tout d'abord nous entrainons deux modèles de langue (ML), l'un correspondant à des données « en-domaine » (ED), l'autre à un sous-ensemble des données « hors-domaine » (HD). Nous donnons ensuite un score $\hat{H}_{Pp}$ à chaque segment c de la partie cible de la table de traduction augmentée (à savoir tbg) avec ces deux modèles de langue (ML_{ED} et ML_{HD}) :

$$\hat{H}_{Pp}(c) = H_{ED}(c) - H_{HD}(c) \tag{2}$$

Ici, $H_{ED}(c)$ est l'entropie croisée (c-à-d le $\log_2$ de la perplexité) de c par rapport à ML_{ED}. $H_{HD}(c)$ est l'entropie croisée de c par rapport à ML_{HD}. Ensuite nous trions l'ensemble des bi-segments d'après leur score ($\hat{H}_{Pp}$) appliqué au côté cible (c). Le score $\hat{H}_{Pp}(c)$, qui peut prendre des valeurs positives ou négatives, est une indication de la « proximité » de la phrase cible c relativement au corpus en-domaine : un score plus bas indique une proximité plus grande [1]. La phase suivante propose de choisir le point de coupure du corpus ainsi trié, grâce à la mesure de perplexité. En ce sens, nous considérons des tranches incrémentales de notre liste triée de bi-segments pour lesquelles nous entrainons un modèle de langue (ML) sur les segments cibles retenus. Enfin, nous calculons la perplexité de chacun de ces modèles de langue sur un corpus de développement en-domaine (ED). Le point de coupure correspond à la perplexité la plus faible ainsi obtenue.

Filtrage bilingue On peut aussi appliquer la procédure précente de façon bilingue. Pour un bi-segment (s, c) , on calcule un score $\hat{H}_{Pp}(s, c)$ de la façon suivante :

$$\hat{H}_{Pp}(s, c) = [H_{ED}(s) - H_{HD}(s)] + [H_{ED}(c) - H_{HD}(c)] \tag{3}$$

Maintenant, le processus de tri est effectué sur le score $\hat{H}_{Pp}(s, c)$, qui dépend des segments source et cible, mais le processus d'identification du point de coupure est effectué seulement sur le côté cible, en nous servant uniquement d'un corpus de développement en-domaine pour la langue cible, comme dans le cas précédent.

4 Expériences

Pour ces expériences, les systèmes de TAS à base de segments ont été entrainés en utilisant l'outil open-source MT Moses (Koehn *et al.*, 2007). Les modèles de langue utilisés sont des n-gram (avec $n = 5$), en appliquant un lissage Kneser-Ney (Chen & Goodman, 1999) grâce aux outils du SRI (Stolcke, 2002). Nous avons utilisé les scores de BLEU [*BiLingual Evaluation Understudy*] (Papineni *et al.*, 2002) et TER [*Translation Edit Rate*] (Snover *et al.*, 2006) comme mesures de performances des modèles pour les expériences.

4.1 Données

Le système de traduction a été entrainé avec les corpus Europarl V.7 (*ep7*) et News-Commentary V.8 (*nc8*), détaillés dans le tableau 3.

1. Ce score est inspiré de (Moore & Lewis, 2010), mais nous l'appliquons pour effectuer un filtrage sur des éléments de la table de bi-segments, alors qu'à l'origine il est appliqué sur des éléments du corpus pour effectuer une sélection au niveau des phrases.

Type	Corpus	# lignes	# mots src (en)	# mots cible (fr)
Apprentissage	*ep7*	2 007 K	56 192 K	61 811 K
	nc8	157 K	4 105 K	4 815 K
Développement	*ntst11*	3 003	75 K	86 K
Évaluation	*tstTerm*	2 577	87 K	103 K

TABLE 3 – Détail des données bilingues utilisées pour les expériences.

Modèle de traduction				*tstTerm (ML référentiel)*		*tstTerm (ML dégradé)*	
Appelation	taille	Nbr. de bi-seg. ajoutés	MHV	BLEU	TER	BLEU	TER
Référentiel	77 203 175	N/A	2 565 (2,9%)	30,7	56,1	27,4	59,0
Base	77 138 148	N/A	5 237 (6,0%)	27,2	58,9	27,2	58,5
Unigrammes	77 138 149	1	2 565 (2,9%)	30,6	56,1	28,4	57,0
Contextes générés (sans filtrage)	78 611 118	1 472 970	2 565 (2,9%)	**31,1**	**55,6**	28,8	56,6
Contextes générés ($\hat{H}_{Pp}(t)$)	77 193 106	54 958	2 565 (2,9%)	30,5	56,1	28,6	57,0
Contextes générés ($\hat{H}_{Pp}(s,t)$)	77 754 665	616 517	2 565 (2,9%)	**31,0**	**55,6**	**29,2**	**56,7**

TABLE 4 – Tableau de statistiques et de résultats pour les différentes configurations appliquées sur le modèle de traductions. Les scores de BLEU et TER sont donnés en pourcentage.

Le corpus de développement est issu de la campagne d'évaluation WMT 2014 (*ntst11*). Le corpus de test spécifique consiste en un ensemble de 2 500 phrases récupérées du corpus MultiUN ($Nations\ Unies$) (Eisele & Chen, 2010) et noté « *tstTerm* ». Ce dernier est donc un bitexte qui contiennent au moins une fois la traduction de « Germany » vers « Allemagne » par phrase, soit 2 672 occurences (environs 3% des mots sources).

4.2 Résultats

Le tableau 4 présente les résultats et les statistiques suivant plusieurs configurations :
– « Référentiel » : un système appris sur les données telles quelles ;
– « Base » : les bi-segments extraits du corpus d'apprentissage sont filtrés pour retirer toute mention à « Germany » et « Allemagne », la table de traduction est entrainée sur les bi-segments restant ;
– « Unigrammes » : la configuration « base » est enrichie par le bi-segment « Germany » → « Allemagne » ;
– « Contextes générés (sans filtrage) » : enrichissement de la configuration « base » par notre approche de génération de contexte (voir section 2) ;
– « Contextes générés ($\hat{H}_{Pp}(s)$) » et « Contextes générés ($\hat{H}_{Pp}(s,t)$) » : respectivement filtrage monolingue et bilingue des contexte générés (description section 3) et ajout des bi-segments du filtrage à la configuration « base » ;
Nous indiquons également la quantité de données générées et les mots hors-vocabulaire (MHV). Avec ces configurations, s'ajoutent deux modèles de langues cibles possibles : un modèle appelé « référentiel », appris sur les données cibles telles quelles, ou alors, un modèle « dégradé » appris sur ces mêmes données mais en supprimant « Allemagne » du vocabulaire. Ceci pour simuler l'absence de ce mot dans le modèle de langue et permettre de mieux voir l'influence du contexte de la table de traduction, indépendamment du modèle de langue.

Les résultats montrent sans surprise une amélioration très significative entre les configurations « Base » et « unigrammes », principalement due à la diminution du nombre de mots-hors-vocabulaire (MHV). On constaste que le système « unigrammes » est du même niveau que le système « Référentiel ». Le système « Contextes générés (sans filtrage) » donne des résultats très encourageants en surpassant le système « Unigrammes » mais également le système « Référentiel ». La contre-performance de ce dernier semble être uniquement lié à des ambiguïtés sur les bi-segments retirés, à savoir toutes les traductions de « Germany » vers « Allemagne ».

Enfin, les deux dernières approches utilisent la sélection monolingue des paires de segments générés avec du contexte (« contexte généré ($\hat{H}_{pp}(t)$) ») et la sélection bilingue de paires de segments (« contexte généré ($\hat{H}_{Pp}(s,t)$) »). La première sélection semble être trop forte : nous observons une diminution des scores BLEU et TER. Cependant, la sélection bilingue de paires de segments nous permet d'être aussi efficace que l'approche « contextes générés (sans filtrage) », mais en ne conservant que 45% des paires de segments générés, et ce, dans les deux tableaux de résultats. Ces dernières expériences semblent valider l'utilisation de la sélection de données bilingues associée à notre approche de génération de contextes. De plus, cette approche permet d'améliorer significativement les résultats par rapport au système « Référen-

tiel ».

5 Etat de l'Art

5.1 Adaptation au domaine en TAS

En Traduction automatique statistique, l'un des sujets les plus étudiés concerne l'adaptation au domaine des systèmes de TAS. Il y a différentes façons d'effectuer cette adaptation. L'une des plus courantes consiste à appliquer une sélection sur les données d'apprentissage. Plusieurs travaux ont été réalisés en utilisant des approches fondées sur la recherche d'information afin d'extraire les parties du corpus qui sont les plus pertinentes (Eck *et al.*, 2004). Des travaux plus récents sont fondés sur l'entropie croisée pour sélectionner les parties les plus pertinentes des données d'apprentissage (Moore & Lewis, 2010; Axelrod *et al.*, 2011).

Dans notre cas, nous nous concentrons sur la traduction en enrichissant le vocabulaire grâce une terminologie spécifique. Or, il n'existe pas de grande quantité de données d'entrainement pour chaque domaine spécifique. C'est pourquoi ces approches ne sont pas adaptées à notre problème. Cependant, ce type d'approche est tout indiqué pour filtrer les bi-segment virtuels générés.

5.2 Enrichissement d'informations lexicales pour la TAS

La plupart des approches dans ce domaine proposent un moyen d'extraire la terminologie spécifique à partir de corpus bilingues (qu'ils soient parallèles ou comparables). Ces approches visent à construire le même genre de dictionnaires que ceux que nous voulons utiliser.

Des travaux antérieurs ont été proposés dans le but de réduire le nombre de mots hors-vocabulaire (MHV) comme (Habash, 2008). Ces approches visent, d'une certaine façon, à ajouter des MHV en les ajoutant dans le corpus d'apprentissage, en utilisant le dictionnaire comme une mémoire de traduction en plus du modèle de traduction. Certains utilisent un pré- ou post-traitement pour éviter le problème MHV (Banerjee *et al.*, 2012; Nikoulina *et al.*, 2012; Tsvetkov *et al.*, 2013). En ce sens, notre approche permet d'éviter tout pré- ou post-traitement lors du processus de traduction et d'utiliser les outils de traduction classique (Koehn *et al.*, 2007) sans modification.

L'approche la plus proche de la nôtre est proposée par (Skadiņš *et al.*, 2013). Dans leur article, ils proposent une technique de sélection de données d'entrainement selon une terminologie spécifique. Cela signifie, qu'ils sélectionnent les bi-phrases qui ne contiennent que la terminologie spécifique recherchée. Ensuite, l'ensemble du processus d'entrainement n'est pas modifié (trouver l'alignement de texte, extraire les bi-segments et enfin estimer les paramètres des modèles de traduction). Enfin, les approches de génération de bi-segments par analogie (Chen *et al.*, 2011; Luo *et al.*, 2013) ne s'intéressent généralement assez peu à la problématique de la terminologie d'un domaine. Or notre approche offre comme principal intérêt d'ajouter la terminologie d'un domaine avec son contexte phrastique. Dans cette catégorie, (Langlais *et al.*, 2009) proposent une approche analogique traduire de la terminologie, cependant, ces hypothèses de traduction ne sont pas intégrées à un système de TAS à base de segments (TAS-BS), contrairement à notre approche.

6 Conclusion et discussion

Cet article présente les premiers résultats sur l'utilisation d'une approche d'enrichissement termnologique automatique pour la traduction automatique statistique. Cette approche comprend principalement l'ajout d'un contexte phrastique associé à cette terminologie. Notre méthode a montré des résultats encourageants en terme de scores de BLEU et TER sur notre corpus de test spécifique. D'un point de vue linguistique, les apports sont principalement situés sur le contexte gauche de la terminologie ainsi insérée. L'un des cas les plus courants et l'ajout d'un déterminant ou d'une préposition correctement traduite.

Prochainement, nous prévoyons d'étendre cette approche à d'autres entités nommées et à d'autres types mots comme les adjectifs, les noms et verbes. Ces derniers sont souvent accompagnés de flexions liées principalement aux accords en genre et en nombre. Enfin, nous souhaitons également confronter notre approche dans le cadre d'une application réelle de post-édition afin de mesurer son impact.

Remerciements

Cette section sera complétée pour la version finale.

Références

AXELROD A., HE X. & GAO J. (2011). Domain Adaptation via Pseudo In-Domain Data Selection. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Edimbourg, Ecosse, Royaume-Uni.

BANERJEE P., NASKAR S. K., ROTURIER J., WAY A. & VAN GENABITH J. (2012). Domain adaptation in SMT of user-generated forum content guided by OOV word reduction : Normalization and/or supplementary data ? In *Proceedings of the 16th Annual Conference of the European Association for Machine Translation (EAMT)*, Trento, Italie.

CHEN B., KUHN R. & FOSTER G. (2011). Semantic smoothing and fabrication of phrase pairs for SMT. In *Proceedings of the International Workshop on Spoken Lanuage Translation (IWSLT-2011)*, San Francisco, Etats-Unis.

CHEN S. F. & GOODMAN J. (1999). An empirical study of smoothing techniques for language modeling. *Computer Speech and Language*, **4**(13), 359–393.

ECK M., VOGEL S. & WAIBEL A. (2004). Language model adaptation for statistical machine translation based on information retrieval. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC)*, Lisbone, Portugal.

EISELE A. & CHEN Y. (2010). MultiUN : A Multilingual Corpus from United Nation Documents. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, Valletta, Malte.

HABASH N. (2008). Four Techniques for Online Handling of Out-of-Vocabulary Words in Arabic-English Statistical Machine Translation. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics : Human Language Technologies (ACL-HLT)*, Colombus, Etats-Unis.

KOEHN P. (2005). Europarl : A parallel corpus for statistical machine translation. In *Proceedings of the Tenth Machine Translation Summit (MT Summit X)*, Phuket, Thaïlande.

KOEHN P., HOANG H., BIRCH A., CALLISON-BURCH C., FEDERICO M., BERTOLDI N., COWAN B., SHEN W., MORAN C., ZENS R., DYER C., BOJAR O., CONSTANTIN A. & HERBST E. (2007). Moses : Open Source Toolkit for Statistical Machine Translation. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*, Prague, République Tchèque.

LANGLAIS P., YVON F. & ZWEIGENBAUM P. (2009). Improvements in analogical learning : Application to translating multi-terms of the medical domain. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, Athenes, Grèce.

LUO J., MAX A. & LEPAGE Y. (2013). Using the productivity of language is rewarding for small data : Populating smt phrase table by analogy. In *Proceedings of the 6th Language & Technology Conference (LTC'13)*, Poznan, Pologne.

MOORE R. C. & LEWIS W. (2010). Intelligent selection of language model training data. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics : Human Language Technologies (ACL-HLT)*, Uppsala, Suède.

NIKOULINA V., SANDOR A. & DYMETMAN M. (2012). Hybrid adaptation of named entity recognition for statistical machine translation. In *Proceedings of the Second Workshop on Applying Machine Learning Techniques to Optimise the Division of Labour in Hybrid MT (ML4HMT - 2012)*, Mumbai, Inde.

PAPINENI K., ROUKOS S., WARD T. & ZHU W.-J. (2002). BLEU : a method for automatic evaluation of machine translation. In *Proceedings of the 41th Annual Meeting of the Association for Computational Linguistics : Human Language Technologies (ACL-HLT)*, Philadelphie, Etats-Unis.

SKADIŅŠ R., PINNIS M., GORNOSTAY T. & VASIĻJEVS A. (2013). Application of online terminology services in statistical machine translation. In *Proceedings of the XIV Machine Translation Summit (MT Summit XIV)*, Nice, France.

SNOVER M., DORR B., SCHWARTZ R., MICCIULLA L. & MAKHOUL J. (2006). A study of translation edit rate with targeted human annotation. In *Proceedings of the 5th Conference of the Association for Machine Translation in the Americas (AMTA)*, Boston, Massachusetts.

STOLCKE A. (2002). SRILM - an extensible language modeling toolkit. In *Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP)*, Denver, Colorado.

TSVETKOV Y., DYER C., LEVIN L. & BHATIA A. (2013). Generating english determiners in phrase-based translation with synthetic translation options. In *Proceedings of the 8th Workshop on Statistical Machine Translation (WMT'13)*, Sofia, Bulgarie.