

# Mesure de confiance de relation sémantique dans le cadre d'un modèle de langage sémantique

Catherine Kobus, Géraldine Damnati et Lionel Delphin-Poulat

France Télécom R&D  
2, avenue Pierre Marzin 22307 Lannion - France

## ABSTRACT

This article proposes a new confidence measure estimated for concept hypotheses given by a semantic language model. This confidence measure is based upon the ontology and the semantic relations linking concepts of a dialog application. It aims at measuring how high a concept hypothesis is related to the other hypotheses of an utterance. The semantic relation confidence measure is evaluated alone and in combination with a classical acoustic confidence measure. It is shown that the two confidence measures are complementary and yield good performance in terms of cross entropy relative reduction.

## 1. INTRODUCTION

Les mesures de confiance sont souvent utilisées en reconnaissance de parole afin de traduire la fiabilité de la réponse (ou hypothèse) fournie par un système de reconnaissance vocale. Les applications des mesures de confiance sont multiples ; elles peuvent être utilisées dans le cadre d'une stratégie de rejet des hypothèses peu fiables ou d'une adaptation du dialogue (demande de confirmation ou non de la phrase prononcée par l'utilisateur), etc. Les mesures de confiance peuvent être évaluées à plusieurs niveaux suivant le but recherché : au niveau de l'hypothèse du mot, du concept ou de la phrase. Il est crucial, dans les applications de reconnaissance vocale et en particulier les applications de dialogue, de déterminer le degré de fiabilité des concepts reconnus.

Les mesures de confiance présentées dans cet article sont calculées sur des hypothèses de concepts, issues d'un modèle de langage sémantique [4], appliqué à un graphe de mots en seconde passe de reconnaissance, dans le cadre d'une application de dialogue. L'application utilisée est liée au domaine bancaire ; elle permet à un utilisateur de consulter ses comptes, ses ordres effectués ou d'effectuer des transactions boursières. Plusieurs méthodes d'estimation de mesures de confiance au niveau des concepts ont déjà été proposées. Dans [3], une mesure basée sur la probabilité *a posteriori* d'un concept est estimée à partir d'un graphe de concepts. [5] introduit une méthode basée sur le consensus de plusieurs classificateurs sémantiques pour valider une hypothèse conceptuelle. [6] incorpore de l'information d'ordre sémantique dans le calcul d'une mesure de confiance sur les concepts, basée sur la probabilité *a posteriori*. La mesure de confiance de relation sémantique présentée dans cet article utilise l'ontologie de l'application de dialogue et, plus précisément, les différentes relations sémantiques pouvant lier les concepts de l'application, ce qui permet de mesurer un indice de cohérence pour un

concept d'une phrase par rapport aux autres concepts de cette même phrase. L'article s'organise de la façon suivante : la première partie définit la mesure de confiance de relation sémantique, et rappelle la définition de la mesure de confiance acoustique. La partie suivante présente la régression logistique utilisée pour calibrer et combiner les mesures de confiance. Enfin sont présentés le contexte expérimental de l'application de dialogue, ainsi que les différents résultats obtenus avec les mesures de confiance (utilisées individuellement ou en combinaison).

## 2. MESURES DE CONFIANCE

### 2.1. Mesure de confiance de relation sémantique

L'idée est d'utiliser la source de connaissance que représentent les relations possibles pouvant exister entre les différents concepts de l'application, afin de vérifier qu'une hypothèse de concept, proposée par le modèle de langage sémantique, est cohérente avec l'ensemble des hypothèses de concepts de la phrase.

**Ontologie de l'application de dialogue** Une ontologie est un ensemble structuré de concepts. Elle décrit comment les différents concepts de l'application sont organisés et liés par des relations sémantiques. L'ontologie pour l'application utilisée est décrite avec le langage KL-One, langage de représentation des connaissances, introduit par Brachman [1]. Ce langage correspond à la première implémentation des réseaux structurés d'héritage ou SIN (pour *Structured Inheritance Network*). Le langage KL-One distingue d'une part les classes conceptuelles, d'autre part les concepts, instances des classes conceptuelles (chaque concept de l'application est une instance d'une seule classe). Il décrit également les relations sémantiques pouvant exister entre certaines de ces classes. Les relations sémantiques sont des relations binaires liant deux classes conceptuelles. Les concepts, instances des classes conceptuelles en relation, sont liés par la même relation. De plus, une relation sémantique liant deux concepts de l'application ne tient pas compte de l'ordre dans lequel ils apparaissent.

Une relation, liant deux classes génériques, est définie de façon unique par la donnée de ces deux classes. Par convention, toute classe conceptuelle est en relation avec elle-même.

La figure 1 représente des exemples de classes conceptuelles en relation dans l'ontologie ; les classes conceptuelles *claNiveau* et *clAIndice* sont liées par la relation *RclaNiveau\_clAIndice*, ainsi que les concepts, instances de chacune de ces classes. Dans le cadre de nos travaux effec-

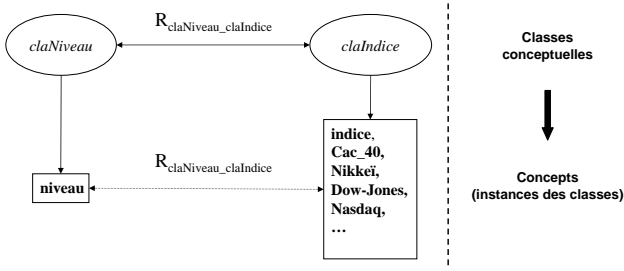


FIG. 1: Exemple de relation décrite par KL-ONE.

tués sur les mesures de confiance, un type de relation particulier a été introduit ; il s'agit de la relation dite "nulle" associée à chaque classe conceptuelle. Un concept, instance d'une classe conceptuelle, est dit en relation nulle lorsqu'il apparaît seul dans un énoncé. Les relations nulles peuvent être vues comme des relations liant chacune des classes conceptuelles à la classe conceptuelle "vide".

**Modèle probabiliste** On veut utiliser les différentes relations sémantiques de l'ontologie afin de définir une distribution de probabilité  $P^{Rel}$ , telle que, pour deux concepts  $c_i$  et  $c_j$  de l'application,  $P^{Rel}(c_i, c_j) = 0$  si les concepts ne sont liés par aucune relation sémantique, et que  $P^{Rel}(c_i, c_j)$  soit fonction de la probabilité de la relation sémantique les reliant, dans le cas contraire. Nous introduisons une fonction  $\gamma$ , qui pour chaque concept  $c_i$ , donne l'unique classe conceptuelle dont il est instance. Chaque concept  $c_i$  étant instance d'une unique classe conceptuelle  $\gamma(c_i)$ , la probabilité de relation  $P^{Rel}(c_i, c_j)$  est définie comme suit :

$$\begin{aligned} P^{Rel}(c_i, c_j) &= P(c_i, c_j | \gamma(c_i), \gamma(c_j)) \cdot P(\gamma(c_i), \gamma(c_j)) \\ &\simeq P(c_i | \gamma(c_i)) \cdot P(c_j | \gamma(c_j)) \\ &\quad \cdot P(\gamma(c_i), \gamma(c_j)) \end{aligned} \quad (1)$$

Nous introduisons alors la relation sémantique de l'ontologie, liant les classes conceptuelles  $\gamma(c_i)$  et  $\gamma(c_j)$  et notée  $R_{\gamma(c_i), \gamma(c_j)}$ . La relation sémantique ne tient pas compte de l'ordre d'apparition des concepts dans la phrase ; on pose ainsi :

$$P(\gamma(c_i), \gamma(c_j)) = P(\gamma(c_j), \gamma(c_i)) = \frac{1}{2} P(R_{\gamma(c_i), \gamma(c_j)}) \quad (2)$$

La probabilité de relation  $P^{Rel}(c_i, c_j)$  devient alors :

$$P^{Rel}(c_i, c_j) = \frac{1}{2} P(c_i | \gamma(c_i)) P(c_j | \gamma(c_j)) P(R_{\gamma(c_i), \gamma(c_j)}) \quad (3)$$

On peut ainsi définir, pour chaque concept  $c_i$  de l'application, une probabilité  $P^{Rel}(c_i)$ , probabilité que le concept  $c_i$  soit en relation avec les autres concepts de l'application. Par définition,

$$P^{Rel}(c_i) = \sum_{c_j} P^{Rel}(c_i, c_j) \quad (4)$$

Les différentes probabilités de relation et d'instanciation des concepts sont estimées par comptage puis lissées à partir d'un corpus d'apprentissage étiqueté en concepts.

**Information mutuelle** Pour définir une mesure de confiance, nous calculons la quantité d'information mutuelle, notée  $i^{Rel}(c_i; c_j)$ , apportée par la co-occurrence des concepts  $c_i$  et  $c_j$  en relation selon la distribution de

probabilité  $P^{Rel}$ .

$$i^{Rel}(c_i; c_j) = \log \frac{P^{Rel}(c_i, c_j)}{P^{Rel}(c_i) \cdot P^{Rel}(c_j)} \quad (5)$$

### Cas d'une hypothèse de phrase contenant au moins deux concepts

Soit  $W$  une hypothèse de phrase, déduite par le modèle de langage sémantique [4] et contenant  $n$  concepts ( $n > 2$ )  $c_1, \dots, c_n$ . Soient  $c_i$  et  $c_j$  deux concepts de cette hypothèse, liés par la relation  $R$  dans l'ontologie. L'information mutuelle  $i^{Rel}(c_i; c_j)$  permet de mesurer l'influence de l'occurrence du concept  $c_i$  sur l'occurrence du concept  $c_j$  dans la même phrase (et inversement). La mesure de confiance de relation sémantique est calculée, pour chaque hypothèse de concept  $c_i$ , en faisant la moyenne de l'ensemble des informations mutuelles de relation du concept  $c_i$  avec les autres concepts de l'hypothèse. L'information mutuelle étant nulle pour deux concepts qui ne sont pas en relation, la moyenne sur l'ensemble des relations possibles *a priori* dans la phrase permet de tenir compte, dans la mesure, des concepts non liés sémantiquement. La mesure de confiance de relation pour un concept  $c_i$  d'une hypothèse constituée de  $n$  concepts ( $n > 2$ ), notée  $MC^{Rel}(c_i)$ , est définie de la façon suivante :

$$MC^{Rel}(c_i) = \frac{1}{n-1} \sum_{j \neq i / Rel(c_i, c_j)=1}^n i^{Rel}(c_i; c_j) \quad (6)$$

La notation  $Rel(c_i, c_j) = 1$  signifie que les concepts  $c_i$  et  $c_j$  sont en relation dans l'ontologie.

### Cas d'une hypothèse constituée d'un seul concept

L'ajout de la relation "nulle" pour chaque classe générique conceptuelle (effective lorsqu'un concept apparaît seul comme hypothèse du modèle de langage sémantique) permet d'étendre la définition de la mesure de confiance au cas d'un concept isolé. L'apprentissage des probabilités sur les relations de l'ontologie permet entre autres de constater les fréquences d'observation de certains concepts suivant les relations. Notamment, certains concepts sont plus fréquemment observés en tant qu'hypothèse unique du modèle de langage sémantique que d'autres. C'est pourquoi une mesure de confiance de relation peut être intéressante dans le cas d'hypothèses de concepts isolés.

Une relation "nulle", notée  $R_\epsilon(X)$ , peut être vue comme une relation liant une classe générique conceptuelle  $X$  à une classe "nulle", noté  $C_\epsilon$ . Le nombre de relations possible étant de 1 pour des hypothèses de concepts isolés, la mesure de confiance de relation  $MC^{Rel}(c)$  associé à un concept  $c$  isolé est égale à l'information mutuelle apportée par l'occurrence d'un concept  $c$  en relation "nulle" :

$$MC^{Rel}(c) = i^{Rel}(c; c_\epsilon) = \log \frac{P^{Rel}(c, c_\epsilon)}{P^{Rel}(c) \cdot P^{Rel}(c_\epsilon)} \quad (7)$$

Dans l'équation 7, la probabilité  $P^{Rel}(C_\epsilon)$  que le concept "nul" soit en relation, est en fait la probabilité d'occurrence de la relation "nulle", égale à la fréquence d'occurrences d'hypothèses de concepts isolés sur le corpus d'apprentissage. La probabilité  $P^{Rel}(c, c_\epsilon)$  peut s'approcher de la façon suivante :

$$P^{Rel}(c, c_\epsilon) \approx P(c | \gamma(c)) \cdot P(R_\epsilon(\gamma(c))) \quad (8)$$

La probabilité  $P(R_\epsilon(\gamma(c)))$  représente la probabilité de la relation "nulle" pour la classe  $\gamma(c)$ .

## 2.2. Mesure de confiance acoustique

Une mesure de confiance acoustique est également utilisée. Cette mesure compare le score de vraisemblance fourni par le modèle acoustique du système de reconnaissance avec le score, qui serait obtenu en utilisant un modèle acoustique simpliste, constitué d'une boucle de phonèmes sans contrainte. Le modèle acoustique du système de reconnaissance vocale est noté  $\mathcal{A}_G$ , tandis que le modèle constitué d'une boucle de phonèmes est noté  $\mathcal{A}_{boucle}$ . Soit  $w$  un mot hypothèse du système de reconnaissance vocale, reconnu avec le modèle acoustique  $\mathcal{A}_G$  et occupant  $T(w)$  trames. Le score de vraisemblance associé au signal acoustique  $X_w$  correspondant à l'hypothèse de mot  $w$  et obtenu avec le modèle acoustique est comparé au score de vraisemblance obtenu pour cette même portion de signal avec le modèle acoustique  $\mathcal{A}_{boucle}$ , constitué d'une boucle de phonèmes sans contrainte. La différence des vraisemblances obtenues est normalisée par le nombre de trames occupées par le mot ; la mesure de confiance acoustique, associée à l'hypothèse de mot  $w$ , notée  $MC^{Acous}(w)$ , est définie comme suit :

$$MC^{Acous}(w) = \frac{[\log(P(X_w|\mathcal{A}_G) - \log(P(X_w|\mathcal{A}_{boucle})))]}{T(w)} \quad (9)$$

La définition de la mesure de confiance est étendue à une hypothèse de concept ([5]). Elle se déduit des mesures de confiance acoustique associées à chaque mot composant le concept. Soit  $c$  une hypothèse de concept, constitué de  $D$  mots  $w_1, w_2, \dots, w_D$  ; la mesure de confiance acoustique associée à l'hypothèse de concept  $c$ , notée  $MC^{Acous}(c)$ , se calcule de la façon suivante :

$$MC^{Acous}(c) = \frac{\sum_{d=1}^D T(w_d) \cdot MC^{Acous}(w_d)}{\sum_{d=1}^D T(w_d)} \quad (10)$$

## 3. UTILISATION DES MESURES DE CONFIANCE

Les mesures de confiance définies précédemment sont calculées pour les différentes hypothèses de concepts issues du modèle de langage sémantique ; le jeu de mesures de confiance associé à chaque hypothèse doit permettre d'estimer la fiabilité, à savoir la probabilité qu'une hypothèse de concept soit correcte étant donné le jeu de mesures de confiance associé.

### 3.1. Critère d'évaluation

Plusieurs méthodes peuvent être utilisées afin d'évaluer des mesures de confiance ; une méthode possible consiste à mesurer la diminution relative de l'entropie croisée d'un corpus de test engendrée par la mesure de confiance. Soit  $C$  un corpus de test constitué de  $N$  hypothèses de concept  $C_i$  ; l'entropie croisée de ce corpus, notée  $H$  est définie de la façon suivante :

$$H = -\frac{1}{N} \left( \sum_{i=1}^N \delta(C_i) \log p_i + (1 - \delta(C_i)) \log(1 - p_i) \right) \quad (11)$$

$\delta(c_i)$  est un indicateur égal à 1 si l'hypothèse de concept  $c_i$  est correcte, égal à 0 sinon. La probabilité  $p_i$  représente la probabilité *a posteriori* que l'hypothèse de concept  $c_i$  soit correcte. Sans l'utilisation de mesure de confiance, cette probabilité *a posteriori* est égale à la précision sur la reconnaissance des concepts, notée *Prec.*, sur le corpus de

test.  $Prec. = \frac{\sum_i \delta(c_i)}{N}$ . On définit ainsi, pour un corpus de test donné, une entropie croisée initiale, notée  $H_{init}$ .

$$H_{init} = -Prec. \log(Prec.) - (1 - Prec.) \log(1 - Prec.) \quad (12)$$

En utilisant un ensemble de mesure de confiance  $MC = mc_1 \dots mc_n$  ( $n \geq 1$ ), la probabilité *a posteriori*  $p_i$  devient  $p_i = P(Cor|MC(c_i))$ , la probabilité  $P(Cor|MC(c_i))$  étant la probabilité qu'une hypothèse de concept  $c_i$  soit correcte étant donné un jeu de mesures de confiance associé  $MC(c_i)$ . Une nouvelle entropie croisée du corpus de test, évaluée avec le jeu de mesures de confiance  $MC$ , notée  $H_{MC}$ , est alors définie.

$$H_{MC} = -\frac{1}{N} \left( \sum_{i=1}^N \delta(c_i) \log(P(Cor|MC(c_i))) + (1 - \delta(c_i)) \log(1 - P(Cor|MC(c_i))) \right) \quad (13)$$

Les mesures de confiance sont ainsi évaluées par la diminution relative induite sur l'entropie croisée d'un corpus de test donné, notée  $\Delta H$  ( $\Delta H = \frac{H_{init} - H_{MC}}{H_{init}}$ ).

### 3.2. Calibration des mesures de confiance

La méthode d'évaluation choisie requiert la connaissance de la probabilité *a posteriori* qu'une hypothèse de concept soit correcte étant donné le jeu de mesures de confiance associé, à savoir la probabilité  $P(Cor|MC(c_i))$ . Dans le cas de l'utilisation d'une seule mesure de confiance  $mc$ , la probabilité à estimer est la probabilité  $P(Cor|mc)$  ; elle est calibrée au préalable sur un corpus de développement et approximée par régression logistique :

$$P(Cor|mc_1) = \frac{1}{1 + e^{-(a_0 + a_1 \cdot mc_1)}} \quad (14)$$

Les paramètres  $a_0$  et  $a_1$  sont estimés de façon à minimiser l'entropie croisée sur un corpus de développement.

### 3.3. Combinaison de mesures de confiance

Plusieurs méthodes peuvent être utilisées afin de combiner des mesures de confiance [2], utilisant des sources de connaissance différentes. Dans le cadre de nos travaux, la régression logistique est utilisée afin de combiner les mesures de confiance acoustique et de relation sémantique. La probabilité *a posteriori* qu'une hypothèse de concept soit correcte étant donné les valeurs des mesures de confiance acoustique et de relation sémantique, notées respectivement  $mc_1$  et  $mc_2$  :

$$P(Cor|mc_1, mc_2) = \frac{1}{1 + e^{-(a_0 + a_1 \cdot mc_1 + a_2 \cdot mc_2)}} \quad (15)$$

Les paramètres  $a_0$ ,  $a_1$  et  $a_2$  sont estimés de façon à minimiser l'entropie croisée du corpus de développement.

## 4. EVALUATIONS

### 4.1. Description des données

Les travaux sur le modèle de langage sémantique et sur les mesures de confiance ont été réalisés dans le cadre d'une application de dialogue [4] dans le domaine bancaire permettant à un utilisateur de consulter ses comptes, d'effectuer diverses transactions ou d'obtenir des informations liées à la bourse. Les différentes probabilités de relation et d'instanciation des concepts, utiles au calcul de la mesure de confiance de relation sémantique, ont été estimées (par

**TAB. 1:** Description des différents corpus de test

	Nb. Conc.	Préc.
<i>Dev_1</i>	1516	12.13
<i>Dev_2+</i>	2956	10.12
<i>Test_1</i>	966	10.15
<i>Test_2+</i>	1894	9.29

comptage puis lissage) sur un corpus d'apprentissage de 24604 énoncés, contenant 34623 concepts. Le modèle de langage sémantique [4], appliqué en seconde passe de reconnaissance à des corpus de test constitués de graphes de mots, donne pour chacun de ces ensembles des hypothèses de concepts. Le premier ensemble, noté *Dev*, constitué de 4472 hypothèses de concepts, est utilisé comme corpus de développement pour la calibration et la combinaison des mesures de confiance. Le deuxième corpus, noté *Test*, est constitué de 2860 hypothèses de concepts et a été utilisé comme corpus de test pour l'évaluation des mesures de confiance. La mesure de confiance de relation sémantique étant définie de façon différente selon le type de concept (isolé ou non dans la phrase), la question se posait alors de savoir s'il fallait fractionner le corpus de développement en conséquence et calibrer différemment les mesures de confiance pour les concepts isolés et non isolés. Les résultats ont montré que le corpus de développement choisi (fractionné ou non) influençait les performances de la mesure de confiance de relation sémantique et qu'il était préférable de fractionner le corpus de développement lors de la calibration de cette mesure de confiance. Les mesures de confiance ont donc été calibrées différemment pour les concepts isolés et les concepts non isolés. Par la suite, les résultats sont présentés en distinguant les hypothèses de phrases des corpus de test contenant une seule hypothèse de celles contenant au moins deux concepts. Ainsi, *Dev\_1* (respectivement *Test\_1*) représente l'ensemble des hypothèses du corpus de développement *Dev* (respectivement du corpus de test *Test*) ne contenant qu'un seul concept, tandis que *Dev\_2+* (respectivement *Test\_2+*) représente l'ensemble des hypothèses du corpus de développement *Dev* (respectivement du corpus de test *Test*) en contenant au moins deux. Le tableau 1 donne pour chaque corpus (de développement et de test) le nombre de concepts et la précision sur la reconnaissance des concepts (et de leur valeur).

#### 4.2. Résultats

Les résultats obtenus en termes de diminution relation relative d'entropie croisée sur les différents corpus de test et avec les différentes mesures de confiance (utilisées seules,  $MC^{Rel}$  et  $MC^{Acous}$  ou en combinaison, *Reg.Log*) sont

**TAB. 2:** Evaluations des mesures de confiance en termes de réduction relative d'entropie croisée ( $\Delta H(\%)$ )

	<i>Dev_1</i>	<i>Dev_2+</i>	<i>Test_1</i>	<i>Test_2+</i>
$MC^{Rel}$	4.06	4.86	2.37	2.99
$MC^{Acous}$	13.11	7.00	10.67	7.43
<i>Reg.Log</i>	17.00	11.58	13.00	10.27

résumés dans le tableau 2. Les résultats montrent que la mesure de confiance de relation sémantique s'avère être aussi efficace pour les hypothèses de concepts isolés que pour les non-isolés, ce qui conforte le choix de l'introduction de relations dites "nulles" dans notre modèle. La mesure de confiance acoustique reste plus performante, sur les différents corpus, que la mesure de confiance de relation sémantique. Elle semble toutefois être moins performante sur les hypothèses de concepts non isolés (à savoir sur les corpus *Dev\_2+* et *Test\_2+*). Une analyse détaillée montre que ces corpus contiennent beaucoup plus de concepts courts en termes de nombre de trames (en moyenne 4 fois plus), par rapport aux corpus de concepts isolés (les corpus *Dev\_1* et *Test\_1*). Une bonne partie de ces concepts courts sont des quantités, pour lesquelles beaucoup d'ambiguïtés peuvent subsister ("*dix*", "*six*", etc). L'analyse acoustique est d'autant plus délicate sur des séquences de signal courtes. C'est pourquoi la mesure de confiance acoustique s'avère moins efficace sur ces corpus.

Les résultats obtenus avec la régression logistique en montrent l'efficacité à combiner des mesures de confiance utilisant des sources de connaissance différentes. Les améliorations apportées individuellement par chaque mesure de confiance sont quasiment cumulées grâce à la régression logistique, ce qui prouve la complémentarité des mesures de confiance.

## 5. CONCLUSION

Cet article propose une nouvelle mesure de confiance de relation sémantique, calculée pour des hypothèses de concept issues d'un modèle de langage sémantique [4]. Cette mesure utilise l'ontologie d'une application de dialogue et les différentes relations sémantiques liant les concepts de cette dernière, afin de mesurer pour chaque hypothèse de concept un degré de cohérence avec les autres hypothèses de concepts de la même phrase. Une mesure est même introduite pour les concepts isolés, seule hypothèse fournie par le modèle de langage sémantique pour une phrase donnée. Les résultats montrent l'intérêt de cette mesure de confiance, les performances étant d'autant plus importantes en combinant, par régression logistique, cette dernière avec une mesure de confiance acoustique.

## RÉFÉRENCES

- [1] T. Brachman. What's in a concept : Structural foundations for semantic networks. *International Journal of Man-Machine Studies*, 9 :127–152, 1977.
- [2] D. Charlet, G. Mercier, and D. Jouvét. On combining confidence measures for improved rejection of incorrect data. In *Eurospeech*, 2001.
- [3] K. Hacioglu and W. Ward. A concept graph based confidence measure. In *ICASSP*, 2002.
- [4] C. Kobus, G. Damnati, L. Delphin-Poulat, and R. De Mori. Conceptual language model design for spoken language understanding. In *Eurospeech*, 2005.
- [5] C. Raymond, F. Béchet, N. Camelin, R. De Mori, and G. Damnati. Semantic interpretation with error correction. In *ICASSP*, 2005.
- [6] R. Sarikaya, Y. Gao, M. Picheny, and H. Erdogan. Semantic confidence measurement for spoken dialog systems. *IEEE Transactions on Speech and Audio Processing*, 2005.