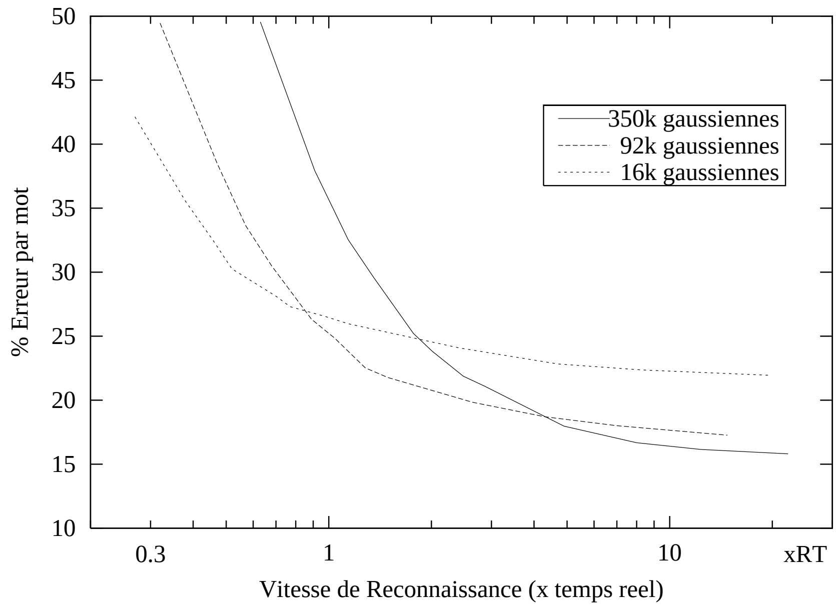

Figure 1. Rapport entre taux d'erreur par mot et vitesse de calcul (sur un Compaq XP1000) pour trois modèles de 350k, 92k et 16k gaussiennes sur un échantillon représentatif d'informations radio- et télédiffusées en anglais américain, avec un modèle de langage 3-gramme sans adaptation du modèle acoustique. Pris ensemble, ces modèles minimisent le taux d'erreur par mot pour une grande plage de vitesse de calcul (de 0,5xTemps Réel à 25xTR).

		Précision moyenne			
données	*méthode*	*base*	*brf*	*pbrf*	*brf+pbrf*
98-R1	*cw*	0.4689	0.5648	0.5591	0.5786
	mw	0.4695	0.5936	0.5574	0.5889
98-S1	*cw*	0.4594	0.5118	0.5621	0.5761
	mw	0.4558	0.5121	0.5884	0.5745
99-R1	*cw*	0.4711	0.5318	0.5147	0.5487
	mw	0.4691	0.5354	0.5098	0.5430
99-S1	*cw*	0.4327	0.5239	0.4919	0.5350
	mw	0.4412	0.5302	0.4943	0.5398

Table1. Comparaison des résultats de recherche d'information sur les données SDR98 et SDR99 en utilisant les pondération Okapi (cw) et markovienne (mw) : recherche de base (*base*), enrichissement de la requête par pseudo-retour de pertinence sur le corpus (*brf*), sur un corpus similaire (*pbrf*) et sur les deux corpus (*brf+pbrf*). R1 : transcription de référence. S1 : transcription automatique.