

Brigitte Bigi - CRCN

Fonction actuelle	Chargée de Recherche CNRS - CRCN
Affectation	<i>Laboratoire Parole et Langage</i> , UMR 7309 Directeur : Laurent Prévot Equipe Interactions
Spécialité et Méthodologie	Traitement automatique des corpus multimodaux
Direction scientifique	<i>INSHS - Institut des Sciences Humaines et Sociales</i>
Délégation régionale	Provence et Corse, DR12
Adresse professionnelle	LPL - 5 avenue Pasteur - 13100 Aix-en-Provence
Téléphone professionnel	04 13 55 27 09
Courrier électronique	brigitte.bigi@cnrs.fr
Page personnelle	http://www.sppas.org/bigi/

Depuis 2017, bénéficiaire de l'obligation d'emploi à cause de capacités très limitées en lecture : RQTH - Reconnaissance de la Qualité de Travailleurs Handicapés attribuée par la commission des droits et de l'autonomie des personnes handicapées.

Études et diplômes

- Thèse de doctorat en informatique (sept. 1997 - sept. 2000), Université d'Avignon.
- DEA informatique et mathématiques, filière Reconnaissance de la parole et logique, Université de la Méditerranée (1997).
- Licence et Maîtrise informatique et mathématiques, Université de la Méditerranée, Marseille (1995-1996).
- DUT Informatique, Université de la Méditerranée, Aix-en-Provence (1994).

Fonctions antérieures

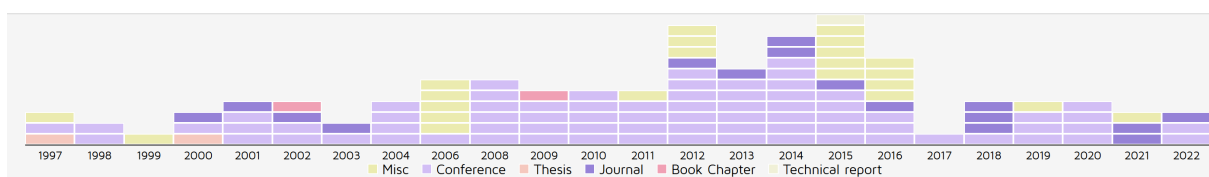
- Chargée de recherche, **LIG** - Laboratoire Informatique de Grenoble, UMR 5217, équipe GETALP. sept. 2002 - juin 2009.
- Chercheuse invitée, **ICSI** - International Computer Science Institute, Berkeley, Californie. janv. 2004 - juil. 2004.
- ATER, **LORIA** - Laboratoire LOrrain de Recherche en Informatique et ses Applications, UMR 7503. oct. 2000 - sept. 2002.
- Doctorante, **LIA** - Laboratoire en Informatique d'Avignon. sept. 1997 - sept. 2000.

Thèmes de recherche récents

Mes travaux, interdisciplinaires par nature, s'inscrivent dans le domaine de la linguistique de corpus (SHS) et celui de la linguistique computationnelle appliquée (Intelligence Artificielle). Dans les projets auxquels je participe, j'apporte mon expertise dans l'analyse automatique de la langue parlée ainsi que dans l'annotation outillée de cette dernière. Je contribue significativement dans les domaines du recueil, de l'annotation automatique, de l'analyse et de la diffusion des grands corpus audio-visuels.

- Formalisation/constitution/extraction de corpus audio-vidéos
- Aide informatisée à l'annotation et l'analyse de corpus multimodaux
- Multilinguisme, langues peu dotées informatiquement

Publications



La liste complète des publications est à l'adresse : <http://www.sppas.org/bigipublications.html>

Articles de revues depuis 2018

B. Bigi, B. Priego-Valverde (2022). *The automatic search for sounding segments of SP-PAS : application to Cheese! corpus*. Human Language Technology - Challenges for Computer Science and Linguistics, Lecture Notes in Artificial Intelligence, Springer, 13212, pp. 16-27.

B. Bigi, A. S. Oyelere, B. Caron (2021). *Resources for Automated Speech Segmentation of the African Language Naija (Nigerian Pidgin)*. Human Language Technology - Challenges for Computer Science and Linguistics, Lecture Notes in Artificial Intelligence, Springer, 12598, pp. 164–173.

C. Meunier, B. Bigi (2021). *Propriétés phonétiques et lexicales dans les variations temporelles des phonèmes en parole conversationnelle*. Studii de lingvistică, 11, pp. 11–38.

B. Bigi (2018). *Annotation representation and file conversion tool*. Contributi del Centro Linceo Interdisciplinare 'Beniamino Segre' (ISSN 0394-0705), 137, pp. 99-116.

B. Bigi, C. Meunier (2018). *Automatic segmentation of spontaneous speech*. Revista de Estudos da Linguagem. International Thematic Issue : Speech Segmentation, 26(4)

B. Priego-Valverde, B. Bigi, S. Attardo, L. Pickering, E. Gironzetti (2018). *Is smiling during humor so obvious? A cross-cultural comparison of smiling behavior in humorous sequences in American English and French interactions*. Intercultural Pragmatics, 15(4), pp. 563–591.

Les 5 publications les plus citées

[466] C. Carpineto, R. De Mori, G. Romano, B. Bigi (2001). *An Information Theoretic Approach to Automatic Query Expansion*, ACM Transactions On Information Systems, 19(1), pp. 1–27.

- [204] B. Bigi (2003). *Using Kullback-Leibler Distance for Text Categorization*, Advances in Information Retrieval : 25th European Conference on IR Research, pp. 305–319, Pisa, Italy.
- [115] B. Bigi (2015). *SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech*, The Phonetician, 111–112, pp. 54–69.
- [107] B. Bigi (2012). *SPPAS : a tool for the phonetic segmentation of speech*, In Proceedings of the Eight International Conference on Language Resources and Evaluation, pp. 1748–1755, Istanbul, Turkey.
- [105] B. Bigi, D. Hirst (2012). *SPeech Phonetization Alignment and Syllabification (SPPAS) : a tool for the automatic analysis of speech prosody*, In Speech Prosody, pp. 19–22, Shanghai China.

Collaborations de ces 5 dernières années

En dehors d'Aix-Marseille Université, ci-après sont listées les universités et autres structures pour lesquelles il y a eu collaboration avec co-publication depuis 2018 :

- Italie** Université de Naples (2022)
- France** Association de parents d'enfants sourds Datha, Paris (2022)
- France** MoDyCo - Modèles, Dynamiques, Corpus, Nanterre (2021)
- Nigéria** Institut Français de Recherche en Afrique - Nigéria (2021)
- Suisse** Université de Lausanne (2020)
- Canada** Université du Québec à Chicoutimi (2020)
- Pologne** Adam Mickiewicz University in Poznań (2019)
- USA** Applied Linguistics Laboratory, Texas A&M University-Commerce, University of Maryland (2018)

Actions de formation

- 2013 - Toulouse** 2h, au CLLE, Univ. Toulouse II Le Mirail.
- 2013 - Aix-en-Pce** 3h, au LPL, adossée au workshop TRASP.
- 2014 - Aix-en-Pce** 2h, au workshop PAC - Phonologie de l'anglais contemporain. LPL.
- 2015 - Barcelone** 2h, Universitat Pompeu Fabra - Master in Applied Linguistics.
- 2015 - Dresde** 3h, Dresde, Allemagne. Tutoriel de la conférence Interspeech.
- 2015 - Hong-Kong** 12h, HK Polytechnic University.
- 2015 - Paris** 8h, Université Paris 8.
- 2015 - Paris** 7h, Consortium IRCOM - Corpus Oraux et Multimodaux, d'HumaNum.
- 2016 - Lyon** 7h, Consortium CORLI - Corpus, Langues et Interactions, d'HumaNum.
- 2016 - Barcelone** 6h, Departament de Filologia Catalana i Lingüística General, Universitat de Barcelona.
- 2017 - Aix-en-Pce** 2h, Formation interne au labo.
- 2018 - Aix-en-Pce** 2h, Formation à l'occasion du TransEVE UNICollaboration : Seminar on research methods to study telecollaboration and virtual exchange, SPHERE Provence.
- 2019 - Paris** 6h, Journées organisées par CORLI - Corpus, Langues et Interactions (HumaNum) sur le thème "Annotation de corpus linguistiques".

Contribution majeure

Depuis 2011, la totalité de mon travail de recherche se présente sous la forme d'un logiciel open source (GPL-v3) et multi-plateforme, SPPAS (<http://www.sppas.org/>). J'ai en effet fait le choix de proposer un outil logiciel regroupant à la fois tous mes travaux de recherche ainsi que les critères nécessaires à son accessibilité à des non-informaticiens, en particulier en répondant aux besoins des chercheurs en linguistique de corpus.



A sa création, SPPAS se destinait essentiellement aux phonéticiens et prosodistes avec 4 annotations automatiques. Aujourd'hui, ce sont 23 annotations qui sont proposées dans le logiciel, ainsi que des outils d'aide à l'analyse. Ainsi, SPPAS peut désormais s'intégrer dans le processus d'annotation de corpus des linguistes de différentes spécialités (phonéticiens, psycholinguistes, neuro-linguistes, etc), comme en témoigne (Ide et al., p. 148) en 2017 : "New developments in providing automatic annotation for linguistic purposes are also appearing, and will lead to the development of new and more efficient workflow practices in this area (e.g.SPPAS [5])."

Les deux indices suivants rendent compte de la pertinence d'avoir mené mon travail de recherche jusqu'au stade de la mise en production :

1. son succès que l'on peut mesurer en terme de téléchargements et de citations de mes travaux. L'utilisation de mes résultats de recherche par un large public international implique nécessairement la robustesse des méthodes proposées, car elles sont employées dans de nombreux contextes différents. Cela constitue une validation scientifique qui va bien au delà de celle d'une publication.
2. le grand nombre de collaborations qui en ont directement découlé. Ce logiciel m'a en effet permis d'établir des collaborations nombreuses et régulières au sein ou en dehors du LPL, dans le cadre de divers projets de recherche, financés ou non. Cet indice est visible via les différents co-auteurs de mes publications.

En 2021, le package de SPPAS a été téléchargé près de 1800 fois par des utilisateurs provenant majoritairement de Chine (23,9 %), France (10,7 %) et États-Unis (9,3 %). Ci-après, la fréquentation du site web de SPPAS ; pour plus de détails, voir : <http://www.sppas.org/analytics/>

	sessions	pages	vues	durée μ session
2016	1746	3145	8 641	3m28s
2017	2238	3847	14 445	5m19s
2018	3259	5527	18 613	5m09s
2019	4016	6692	19 499	4m33s
2020	3531	5764	16 753	4m36s
2021	4593	7045	18 488	3m30s

En 5 ans, j'ai entièrement développé, documenté et diffusé 33 releases : SPPAS est passé de la version 1.9.4 en janvier 2018 à la version 4.6 en août 2022.

Fiche synthétique :

Nom	SPPAS - the automatic annotation and analysis of speech
Site web	http://www.sppas.org
Forge	http://sppas.sf.net
Software Management Plan	https://dmp.opidor.fr/plans/4296/export.pdf
Software Heritage Identifier	swh :1 :rev :6f09ed09165ac185912f4c8603d0bc9e63f398e3
Documentation logiciel	https://hal.archives-ouvertes.fr/hal-01414600v7
Documentation ressources	https://hdl.handle.net/11403/sppasresources
Documentation format natif	https://hal.archives-ouvertes.fr/hal-03468454

Prix

SPPAS fait parti des 7 lauréats que compte le CNRS parmi les 10 **prix science ouverte du logiciel libre de la recherche**. Il se distingue par le fait qu'il est le seul à relever des Sciences Humaines et Sociales et qu'il est le seul à être produit par une seule personne et non par une équipe.

Les prix science ouverte du logiciel libre de la recherche ont été mis en place pour la première fois en 2022. Ils ont été impulsés par le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation (MESRI) et portés par le comité pour la science ouverte (<https://www.ouvrirelascience.fr>). La création de ce prix s'inscrit dans cadre du deuxième Plan national pour la science ouverte.

Les prix ont été décernés durant la conférence OSEC (Paris Open Science European Conference) en février 2022. Ils ont été mis en perspective par une session consacrée au logiciel comme pilier de la science ouverte. C'est à cette occasion qu'a été élaboré l'appel de Paris sur l'évaluation de la recherche (<https://osec2022.eu/fr/appel-de-paris/>). Cet appel indique explicitement le fait que *“le système actuel d'évaluation de la recherche, des chercheurs et des institutions de recherche n'encourage et ne récompense pas assez la qualité de l'ensemble des productions de la recherche dans toute sa diversité”*.

La création de ce prix se donne pour objectif de mettre en valeur les projets et les équipes de recherche qui œuvrent au développement et à la diffusion des logiciels libres et qui contribuent à la construction d'un bien commun de première importance. Leur objectif est de **reconnaître la contribution importante que constitue la production de logiciels libres** et d'attirer l'attention de la communauté scientifique sur des réalisations d'exception ou très prometteuses, qui peuvent servir de modèle pour les prochaines générations de chercheurs et d'ingénieurs. Les prix ont été déclinés en trois catégories, qui récompensent :

- la qualité scientifique et technique du logiciel
- la construction d'une communauté active de contributeurs et utilisateurs
- l'effort indispensable de fournir une documentation qui facilite l'usage et l'appropriation des logiciels.

De plus, le Prix du jury récompense un projet exemplaire rassemblant les composantes techniques, l'animation de communauté et la documentation. Le jury a également décerné des accessits dans chacune des catégories.

SPPAS a été récompensé par un Accessit au Prix spécial du Jury.

Le prix s'est accompagné d'un trophé décerné par Mme la ministre Sylvie Retailleau.