

# Lim Lian Tze, Ph.D.

 [liantze@gmail.com](mailto:liantze@gmail.com)    [@liantze](https://twitter.com/liantze)    <http://liantze.penguinattack.org/>  
 [orcid.org/0000-0003-0796-8762](https://orcid.org/0000-0003-0796-8762)    <http://www.linkedin.com/in/limliantze/>

## Employment History

- 2014 – ....  **Community TeXpert.** Overleaf.com.
- 2013 – 2015  **Lecturer.** Information Technology Department, School of Engineering, Science and Technology, KDU College Penang, Malaysia. (Part-time position Sep – Dec 2015.)
- 2013 – 2014  **Programme Leader.** Information Technology Department, School of Engineering, Science and Technology, KDU College Penang, Malaysia.
- 2011 – ....  **Freelance L<sup>A</sup>T<sub>E</sub>X Trainer and Consultant.** Clients include the Linux Foundation; Al-Andalus Legal Consulting; Consumers International; NAV6, Universiti Sains Malaysia; USAINS Sdn Bhd; Universiti Malaysia Sarawak; Universiti Utara Malaysia; Universiti Kebangsaan Malaysia; Universiti Malaysia Perlis; Universiti Malaya; University of Southern Australia; Nanyang Technological University.
- 2003 – 2009  **Research Officer.** Computer Aided Translation Unit, School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia.
- 2002 – 2003  **Tutor/Assistant Lecturer.** Dept. of Information Technology, College of Engineering, Universiti Tenaga Nasional, Selangor, Malaysia.
- 2001 – 2002  **Trainee Executive.** Tenaga Nasional Berhad, Kuala Lumpur, Malaysia.
- 2000  **Open Day Guide.** Department of Computer Science, University of Warwick, United Kingdom.
- 1999 – 2000  **IT Trainer.** University of Warwick Student Union, United Kingdom.

## Education

- 2009 – 2013  **Ph.D., Multimedia University, Malaysia** in Natural Language Processing.  
Thesis title: *Low-Cost Multilingual Lexicon Construction for Under-Resourced Languages*. More details at <http://liantze.penguinattack.org/phd.html>
- 2003 – 2006  **M.Sc. Computer Science, Universiti Sains Malaysia, Malaysia** in Automated Translation.  
Thesis title: *Using Conceptual Vectors to Improve Translation Selection*.
- 1998 – 2001  **B.Sc. (Hons) Computer Science, University of Warwick, United Kingdom.**  
*First Class Honours.* Department Prize for Outstanding Performance (2nd place in class).
- 1996 – 1998  **A-Levels, Kolej Yayasan Pelajaran MARA Kuala Lumpur, Malaysia.**  
Physics (A), Advanced Pure Mathematics (A), Advanced Mechanics and Applied Mathematics (A).
- 1991 – 1995  **SPM, Sekolah Menengah Sin Min, Sungai Petani, Kedah, Malaysia.**  
A1: English, Bahasa Malaysia, Bahasa Cina, Mathematics, Additional Mathematics, Physics, Chemistry, Biology, History, Cambridge English 1119. A2: Moral Education.  
(Recipient of a full TNB scholarship award for A-levels and undergraduate studies.)

## Teaching

- 2015  **Introduction to Natural Language Processing** open seminar to KDU staff and students
- 2014 – 2015  **Natural Language Processing and WordNet** seminars at Kasetsart University; Nanyang Technological University; Dewan Bahasa dan Pustaka
- 2013 – 2015  **Final Year Project Tutor/Coordinator, B. Sc.(Hons) Computer Studies**  
 **Visual Basic.NET Programming** lecturer, Dip. Comp. Studies.
- 2013  **Artificial Intelligence for Games** lecturer, Dip. Games Tech.
- 2013 – 2014  **Internet Technology & Applications (HTML/CSS, XML, C#.NET, ADO.NET)** lecturer, Dip. Comp. Studies and Dip. Games Tech.
- 2010 – ....  **L<sup>A</sup>T<sub>E</sub>X workshops** at various universities.
- 2008  **Lexical Semantics** assisting in lectures for Natural Language Processing, B. Sc.
- 2003  **Lotus Notes and LotusScript** lab tutor for Management Information Systems, B. IT.
-  **DHTML/CSS, ASP, JavaScript** lab tutor for Cyber Technology, D. IT.
- 2002  **Office productivity suite and Visual Basic** lab tutor for Computing Skills, B. IT.
-  **SQL** lab tutor for Database Systems, D. IT.
- 1999 – 2000  **European Computer Driving License (ECDL)** trainer.

## Research Interests

- Multilingual lexicons and machine translation
- Word sense disambiguation and translation selection
- Programming languages and paradigms
- Computer typesetting with L<sup>A</sup>T<sub>E</sub>X and XML/XSLT tools

## Research Publications

- 1 Lim, L. T., Chiew, R. T., Tang, E. K., Rusli, A. G., & Naimah, Y. (2016). Digitising a machine-tractable version of Kamus Dewan with TEI-p5. *PeerJ Preprints*, 4, e2205v1. doi:10.7287/peerj.preprints.2205v1
- 2 Bond, F., Lim, L. T., Tang, E. K., & Riza, H. (2014). The combined Wordnet Bahasa. *NUSA: Linguistic studies of languages in and around Indonesia*, 57, 83–100. doi:10.7287/peerj.preprints.2205v1
- 3 Boon, K. M. & Lim, L. T. (2014). An examination question paper preparation system with content-style separation and Bloom's taxonomy categorisation. In *Proceedings of the 3rd International Conference on E-Learning and E-Technologies in Education (ICEEE 2014)* (pp. 39–47). Kuala Lumpur, Malaysia. doi:10.7287/peerj.preprints.2205v1
- 4 Lim, L. T., Soon, L.-K., Lim, T. Y., Tang, E. K., & Ranaivo-Malançon, B. (2014). Lexicon+TX: rapid construction of a multilingual lexicon with under-resourced languages. *Language Resources and Evaluation*, 48(3), 479–492. doi:10.1007/s10579-013-9253-0
- 5 Lim, L. T. (2013). *Low-cost multilingual lexicon construction for under-resourced languages* (Doctoral dissertation, Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia).
- 6 Lim, L. T., Soon, L.-K., Lim, T. Y., Ranaivo-Malançon, B., & Tang, E. K. (2013). Context-dependent multilingual lexical lookup for under-resourced languages. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (pp. 294–299). Sofia, Bulgaria.
- 7 Lim, L. T., Ranaivo-Malançon, B., & Tang, E. K. (2011a). Low cost construction of a multilingual lexicon from bilingual lists. *Polibits*, 43, 45–51.
- 8 Lim, L. T., Ranaivo-Malançon, B., & Tang, E. K. (2011b). Symbiosis between a multilingual lexicon and translation example banks. *Procedia: Social and Behavioral Sciences*, 27, 61–69.
- 9 Lim, L. T., Ranaivo-Malançon, B., & Tang, E. K. (2011c). Symbiosis between a multilingual lexicon and translation example banks. In *Proceedings of the 12th Conference of the Pacific Association for Computational Linguistics (PACLING 2011)*. Kuala Lumpur, Malaysia.
- 10 Belaton, B. & Lim, L. T. (Eds.). (2009). *Grid computing cluster: the development and integration of grid services and applications*. Penang, Malaysia: Platform for Information & Communication Technology Research, Universiti Sains Malaysia.
- 11 Lim, L. T. (2009). Multilingual lexicons for machine translation. In *Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services (iiWAS2009) Master and Doctoral Colloquium (MDC)* (pp. 732–736). Kuala Lumpur, Malaysia.
- 12 Tang, E. K., Lim, L. T., Ye, H. H., & Schwab, D. (2009). Grid-enabled Blexisma2. In B. Belaton & L. T. Lim (Eds.), *Grid computing cluster: the development and integration of grid applications and services* (pp. 23–26). Penang, Malaysia: Platform for Information & Communication Technology Research, Universiti Sains Malaysia.
- 13 Lim, L. T. & Schwab, D. (2008). Limits of lexical semantic relatedness with ontology-based conceptual vectors. In *Proceedings of the 5th International Workshop on Natural Language Processing and Cognitive Science (NLPSC'08)* (pp. 153–158). Barcelona, Spain.
- 14 Schwab, D. & Lim, L. T. (2008). Blexisma2: a distributed agent framework for constructing a semantic lexical database based on conceptual vectors. In *Proceedings of the 1st International Conference on Distributed Frameworks & Applications 2008 (DFMA 2008)* (pp. 102–110). Penang, Malaysia.
- 15 Lim, L. T., Tan, E. H., & Tang, E. K. (2007). Digitising dictionaries for advanced look-up and lexical knowledge research in Malay. In *Proceedings of the 11th International Conference on Translation*. Kuala Lumpur, Malaysia.
- 16 Schwab, D., Lim, L. T., & Lafourcade, M. (2007a). Conceptual vectors, a complementary tool to lexical networks. In *Proceedings of the 4th International Workshop on Natural Language Processing and Cognitive Science (NLPSC 2007)*. Funchal, Madeira, Portugal.
- 17 Schwab, D., Lim, L. T., & Lafourcade, M. (2007b). *Les Vecteurs Conceptuels, un Outil Complémentaire aux Réseaux Lexicaux*. In *Proceedings of Traitement Automatique des Langues Naturelle (TALN 2007)*. Toulouse, France.
- 18 Lim, L. T. (2006a). Improving translation selection with conceptual vectors. In *Proceedings of the Regional Computer Science Postgraduate Colloquium (RecSPC'06)* (pp. 231–234). Penang, Malaysia.
- 19 Lim, L. T. (2006b). *Improving translation selection with conceptual vectors* (MSc thesis, School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia).

- 20 Lim, L. T. & Hussein, N. (2006). Fast prototyping of a Malay WordNet system. In *Proceedings of the Language, Artificial Intelligence and Computer Science for Natural Language Processing (LAICS-NLP) Summer School Workshop* (pp. 13–16). Best Paper Award. Bangkok, Thailand.
- 21 Lim, L. T., Tang, E. K., & Zaharin, Y. (2005). A formal test specification translator using ontology and XML/XSLT technologies. In *Proceedings of the 1st Malaysian Software Engineering Conference (MySEC'05)* (pp. 155–160). Best Paper Award. Penang, Malaysia.
- 22 Lim, L. T. & Tang, E. K. (2004). Building an ontology-based multilingual lexicon for word sense disambiguation in machine translation. In *Proceedings of the 5th Workshop on Multilingual Lexical Database (Papillon 2004)*. Grenoble, France.

## R&D Projects

- **WordNet Bahasa:** in collaboration with Nanyang Technological University, University of Indonesia, etc (<http://wn-msa.sourceforge.net/>)
- **Annotating Kamus Dewan Micro- and Macro-Structures for Enhanced Indexing and Search:** collaboration project with The Name Technology Sdn. Bhd. under MDEC's Innovation Voucher scheme
- **Lexicon+TX:** (Ph.D. project) a flexible, low-cost multilingual lexicon generation framework, especially suitable for under-resourced languages
- **Semantic Lexical Base:** using conceptual vectors and lexical functions to model lexical knowledge
- **Blexismaz2:** a distributed agent framework for constructing a Semantic Lexical Base
- **Multilingual Lexiconontology:** designing and constructing an ontology-based multilingual sense dictionary and a sense-tagger based on the generated semantic data
- **Malay WordNet prototype** from existing bilingual dictionary resources.
- **Word sense disambiguation with conceptual vectors** using dictionaries, ontologies and translation corpus as knowledge sources.
- **Translation selection** for an example-based machine translation system.
- **Development of multilingual machine-readable dictionaries** from paper dictionaries, to cater for NLP purposes as well as look-ups by human users.
- **Formal Tester Specification Translator** which translates test programs between different platforms, using XML/XSLT and ontology technologies.
- **L<sup>A</sup>T<sub>E</sub>X style and template files** for USM, MMU, UM, UUM, UKM, Unimas and Caltech theses.

## Skills

- |           |   |
|-----------|---|
| Languages | ■ Strong reading, writing and speaking competencies for English, Mandarin Chinese, Malay. Some knowledge of Japanese, French and Spanish. |
| Coding    | ■ Java, PHP, Python, R, SQL, XML/XSL, L <sup>A</sup> T <sub>E</sub> X, ASP.NET, Visual Basic.NET, C#.NET, LotusScript, Prolog.            |
| Databases | ■ MySQL, PostgreSQL, HSQL, SQLite, LotusNotes.  |
| Web Dev   | ■ HTML, CSS, JavaScript, Apache Web Server, Tomcat Web Server.  |
| Misc.     | ■ Academic research, teaching, training, consultation, L <sup>A</sup> T <sub>E</sub> X typesetting and publishing.                        |

## Miscellaneous Experience

### Awards and Achievements

- |      |   |
|------|---|
| 2015 | ■ <b>Certificate of Achievement in R Programming</b> , Coursera online course by Johns Hopkins University.  |
| 2014 | ■ <b>Certificate of Achievement in Mining Massive Datasets</b> , Coursera online course by Stanford University.   |
|      | ■ <b>Certificate of Achievement in the Data Scientist's Toolbox</b> , Coursera online course by Johns Hopkins University.   |
| 2012 | ■ <b>Certificate of Achievement in Natural Language Processing</b> , Coursera online course by Stanford University.   |
|      | ■ <b>Certificate of Achievement in Machine Learning</b> , Coursera online course by Stanford University.  |
|      | ■ <b>Certificate of Achievement in Computing for Data Analysis</b> , Coursera online course by Johns Hopkins University.  |
| 2006 | ■ <b>Best Paper Award</b> , Language, Artificial Intelligence and Computer Science for Natural Language Processing (LAICS-NLP) Summer School Workshop, Bangkok, Thailand. |
| 2005 | ■ <b>Best Paper Award</b> , First Malaysian Software Engineering Conference (MySEC'06), Penang, Malaysia.   |
| 2002 | ■ <b>Merit Award</b> , 25-day Standard Course SC450, Outward Bound School, Lumut, Malaysia.   |

## Miscellaneous Experience (continued)

- 2001 ■ Department Prize for Outstanding Student Performance, Department of Computer Science, University of Warwick, United Kingdom.

## Certification

- 2014 ■ TRIZ Level 1 Practitioner. Awarded by MyTRIZ Innovation Association, Malaysia.  
2006 ■ Certified IT Professional for FITPE (Fundamental IT Professional Examination). Awarded by Ministry of Higher Education Malaysia.

## Service

- 2014 ■ Reviewer, The 28th Pacific Asia Conference on Language, Information and Computing (PACLIC28), Phuket, Thailand.  
■ Trainer in Outcome-based Teaching & Learning (Delivery Skills), KDU College Penang.  
2014 – 2015 ■ Research & Development Committee Member, KDU College Penang.  
2013 – 2015 ■ Teaching & Learning Committee Member, KDU College Penang.  
2012 ■ Keynote Speaker, The 6th International Workshop on Malay and Indonesian Language Engineering (MALINDO) 2012.  
2012 – 2014 ■ Trainer, *L<sup>A</sup>T<sub>E</sub>X Workshops* at various institutions.  
2011 ■ Invited Speaker, *L<sup>A</sup>T<sub>E</sub>X: More Than Just Academic Papers and Theses*, the Malaysia Open Source Conference 2011.  
2010 ■ Website maintenance team member, Natural Language Processing Special Interest Group, Faculty of Information Technology, Multimedia University.  
2009 ■ Editor and L<sup>A</sup>T<sub>E</sub>X designer/typesetter, *Grid Computing Cluster: The Development and Integration of Grid Services and Applications* (Grid@USM 2008/09 Report).  
■ Reviewer, The Second International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2009), London, United Kingdom.  
2008 ■ Session Chair, the 1st International Conference on Distributed Frameworks & Applications 2008 (DFMA 2008), Penang, Malaysia.  
2007 ■ Session Chair, the 11th International Conference on Translation, Kuala Lumpur, Malaysia.  
■ Workshop Speaker, *Introduction to L<sup>A</sup>T<sub>E</sub>X*, Computer Sciences Postgraduate Colloquium 2007, School of Computer Sciences, Universiti Sains Malaysia.  
■ L<sup>A</sup>T<sub>E</sub>X Typesetter, Programme/Abstract Booklet of Computer Sciences Postgraduate Colloquium 2007, School of Computer Sciences, Universiti Sains Malaysia.  
2004 ■ Organising Committee, Computer Sciences Postgraduate Colloquium, School of Computer Sciences, Universiti Sains Malaysia.  
2002 ■ Facilitator and Speaker, Kem Remaja Bestari 2002, Tenaga Nasional Berhad, Malaysia.

## References

Available on Request