José David Lopes

KTH Speech, Music and Hearing
Lindstedsvägen 24
10044 Stockholm
Sweden

jdlopes@kth.se

1 Research Interests

The end goal of my research is to improve communication between humans and between humans and machines. Communication may fail due to a number of reasons: noisy channel, lack of common ground, information about context or even difficulty for one to speak in a different language. While solving all communication problems might be an utopia, my aim is to address smaller problems inherent to one of the reasons above mentioned. For this, I am interested in studying **human-human dialogues** and then apply the outcome of these studies to Spoken Dialogue Systems (SDSs). These studies may show how humans establish common ground and how they handle errors in communication. Therefore they can contribute to the development **SDSs with human-like behavior**. The language barrier is yet another problem that adds to the communication breakdown pile. One of the my long term goals is to create an **SDSs to improve conversation skills** in new language.

1.1 Previous Work

Establishing a common ground in dialogues encompasses several dimensions. During my PhD I have addressed the problem of finding a common ground with respect to lexical items in a dialogue system. The goal was to find a common words that were accepted by both the system and the user to improve the performance of the system. There were two motivations for this. The first one is that there are words that are more easily recognized by a speech recognizer than others. The second one was that there are words that may be preferred by some users and thus the system should try to use them to increase the degree of engagement with the users. To achieve this goal we first studied whether there were words/expressions that users picked up more easily than others (Lopes et al., 2011). In the following step we tried to improve the confidence measure provided by the system to know when a word/expression proposed was not accepted (Lopes et al., 2012). We have built our first rule-based model to automatically choose the ideal word/expression to be used in the system prompt (Lopes et al., 2013). And finally we have also explored a data-driven method to do the same task (Lopes et al., 2015a).

Since I moved to KTH, I started to investigate how to automatically detect communication breakdowns in dialogues with SDSs. Repetitions could be a sign that the communication between the system and the user might have been broken. An automatic method to detect repetitions in interactions with SDSs was developed (Lopes et al., 2015b). This method could be applied to a more general framework to detect miscommunications in dialogues such as the one presented in (Meena et al., 2015). But ultimately, if the goal is to detect the source of the detected miscommunications need to be filtered out by a method such as the one presented in (Georgiladakis et al., 2016). The datasets used in this studies are publicly available (Lopes et al., 2016).

2 Future of Spoken Dialog Research

Where do you think the field of dialogue research will be in 5 to 10 years? The advances in the last few years were rather little, which makes me think that the next coming years will be very exciting. It seems that current solutions for rule-based and data-driven dialogue management have reached their limits. New statistical methods that will be combined with knowledge resources to provide robust and long-term interaction between users and systems.

What do you think this generation of young researchers could accomplish in that time? These generation of researchers should take the opportunity that the field is under the spotlight right now to develop methods that can make spoken dialogue systems something that people use their everyday life.

What kind of questions need to be investigated to get the field to that point? In my opinion, the leading researchers in the field will have to bring new statistical methods into the field that can incorporate information that is today provided either by rule-based models or ad-hoc handcrafted rules.

3 Suggestions for Discussion

- Experiment design and evaluation methods for non-task oriented dialogues and for SDSs for long term interactions.

- Human likeliness in SDSs: Should we aim for systems that behave like humans? Should we use human behavior as a role model that guides the systems we develop? Or should we aim for a different kind of behavior that does not frustrate the user expectations?

- Impact on society: what would be the consequences of the systems will be to the society, what kind of systems should we aim to build, should there be any sort of regulation to limit our research.

References

S. Georgiladakis, G. Athanasopoulou, R. Meena, J. Lopes, A. Chorianopoulou, E. Palogiannidi, E. Iosif, G. Skantze, and A. Potamianos. 2016. Root cause analysis of miscommunication hotspots in spoken dialogue systems. In *Interspeech 2016*, San Francisco, CA, USA, sep.

José Lopes, Maxine Eskenazi, and Isabel Trancoso. 2011. Towards choosing better primes for spoken dialog systems. In *2011 IEEE Workshop on Automatic Speech Recognition & Understanding, ASRU 2011, Waikoloa, HI, USA, December 11-15, 2011*, pages 306–311.

José Lopes, Maxine Eskenazi, and Isabel Trancoso. 2012. Incorporating ASR information in spoken dialog system confidence score. In *Computational Processing of the Portuguese Language - 10th International Conference, PROPOR 2012, Coimbra, Portugal, April 17-20, 2012. Proceedings*, pages 403–408.

José Lopes, Maxine Eskenazi, and Isabel Trancoso. 2013. Automated two-way entrainment to improve spoken dialog system performance. In *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2013, Vancouver, BC, Canada, May 26-31, 2013*, pages 8372–8376.

José Lopes, Maxine Eskenazi, and Isabel Trancoso. 2015a. From rule-based to data-driven lexical entrainment models in spoken dialog systems. *Computer Speech & Language*, 31(1):87–112.

José Lopes, Giampiero Salvi, Gabriel Skantze, Alberto Abad, Joakim Gustafson, Fernando Baptista, Raveesh Meena, and Isabel Trancoso. 2015b. Detecting repetitions in spoken dialogue systems using phonetic distances. In *Interspeech 2015*.

J. Lopes, A. Chorianopoulou, E. Palogiannidi, H. Moniz, A. Abad, K. Louka, E. Iosif, and A. Potamianos. 2016. The spedial datasets: datasets for spoken dialogue systems analytics. In *10th edition of the Language Resources and Evaluation Conference*, may.

R. Meena, J. Lopes, G. Skantze, and J. Gustafson. 2015. Automatic detection of miscommunication in spoken dialogue systems. In *Proceedings of 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 354–363, Prague, Czech Republic, sep. Association for Computational Linguistics.

Biographical Sketch

José David Lopes is currently a post-doc at the KTH Speech, Music and Hearing department. He got his master from the University of Coimbra in 2008, working on acoustic modelling for noise adverse environments. In the same he joined the Spoken Language System Lab at INESC-ID in Lisbon to start his PhD, which he received form the Technical University of Lisbon in 2013. His thesis was entitled "Lexical entrainment in Spoken Dialogue Systems". After completing his PhD and before starting as a Post-doc at KTH, he worked at the Speech Interactive Research Group at University of the Basque Country, developing a Dialogue System for bus schedule information in the city of Bilbao. After working on the SpeDial european project to develop tools for spoken dialogue system analytics, José's current research is on detecting and developing pauses for Spoken Dialogue Systems.