Bringing Pragmatics to Porttinari – Adding Speech Acts to News Texts

Nataly L. P. da Silva
EACH/USP
São Paulo – SP – Brazil
natalypatti@usp.br

Norton T. Roman
EACH/USP
São Paulo – SP – Brazil
norton@usp.br

Ariani Di Felippo
DLL/UFSCar
São Carlos - SP - Brazil
ariani@ufscar.br

Abstract

The automatic classification of Speech Acts is a topic of great interest within the NLP area which could be taken as a first step towards the semantic representation of texts. However, in order to carry out this task, a reasonable amount of annotated data is necessary, if one is to apply any Machine Learning (ML) technique to this end. In this work, we present a subset of the Porttinari-base corpus manually annotated with Speech Acts, following an adapted version of the ISO-24617-2 standard, so as to provide the community with a starting point for automatic identification of speech acts in news texts written in (Brazilian) Portuguese. To illustrate the corpus' usefulness, we also present the results of training an ML distributional model to classify speech acts in such texts.

1 Introduction

Speech Acts theory (Austin, 1962) states that when we say something we not only communicate the (composite) semantic content of the words we pronounce, but also execute an action in doing so. Within this setting, a speech act represents our communicative intention when we express ourselves. That is through language it is possible to perform an action or have someone perform an action such as thanking, questioning, asking, promising etc.

The automatic classification of speech acts consists of associating classes of speech acts (*e.g.* asking, stating, promising etc.) to each utterance in certain contexts, such as lines in a dialog or tweets or sentences in a text, with the objective of identifying the communicative function performed by that utterance. This could be taken as a first step towards the semantic representation of texts, as language understanding involves the ability to relate text structure to the world and understanding the communicative intention of speech (Bender and Koller, 2020). To do so, among other things one needs corpora annotated with Speech Act classes.

Although such corpora can be found for different application areas, languages and using different taxonomies (*e.g.* MapTask (Thompson et al., 1993), Discourse Annotation and Markup System of Labeling (DAMSL) (Core and Allen, 2001)), there seems to be a lack of corpora regarding speech acts in Portuguese, specially for news texts. To help fill in this gap, in this article we present a subset of the Porttinari-base corpus[1], manually annotated with Speech Acts according to the ISO 24617-2 taxonomy (ISO, 2012).

Among our main contributions, we built one of the few (if not the first) corpus of news texts in Portuguese annotated with Speech Acts, adding new resources to the Porttinari (PorTuguese Treebank) project. Additionally, we present an adaption of the ISO 24617-2 taxonomy, so it can be applied to news texts but without loosing its role as a standard. As a final contribution, we trained an ML distributional model – BERTimbau (Souza et al., 2020) – to automatically classify speech acts in news texts annotated with this taxonomy.

The annotated corpus is available[2] to the community, under a Creative Commons license. We expect it, along with the preliminary results on applying BERTimbau to this task, to serve both as a resource and baseline to other researchers in the area. The rest of this paper is organised as follows. Section 2 gives an overview of some related initiatives on speech acts annotation. Next, in Section 3, we describe in more detail the Porttinari-base corpus, along with the ISO 24617-2 taxonomy, the annotation procedure and experimental setup we followed. Our results are presented and discussed in Section 4, while in Section 5 we present our final remarks and directions for future work.

[1] https://sites.google.com/icmc.usp.br/poetisa/porttinari

[2] https://github.com/natalypatti/porttinari-base-speech-acts

2 Related Work

There are several data sets currently available for Speech Acts classification, such as SwDA (Jurafsky and Shriberg, 1997), MRDA (Shriberg et al., 2004) and MAPTASK (Thompson et al., 1993), which are mainly focused either on free dialogues between two or more parties or on task-oriented situations. Much of the extant work, however, opts for building its own data sets (*e.g.* (Chen and Di Eugenio, 2013; Blache et al., 2020)), usually tailored to a specific problem, which cannot be addressed with currently available corpora. This, in turn, highlights the need for more diverse data sets to be built, covering different genres, domains and styles, so as to speed up future research, saving it from this laborious task.

In the annotation of these corpora, different tagsets of speech acts are employed. In SwDA, for example, 1.115 conversations from the Switchboard corpus (Godfrey et al., 1992) were annotated according to the SwDA-DAMSL taxonomy, which comprises 42 tags. MRDA, in turn, defines a hierarchical taxonomy, based on the SwDA-DAMSL classes, thereby allowing researchers to focus on the highest level classes, with only 5 speech acts. Finally, MAPTASK delivers a task-oriented corpus with 13 speech acts tailored to a specific area of interest.

This variability of annotation schemes leads to some negative aspects related to standardisation, reuse and comparison. In this regard, ISO 24617-2 (ISO, 2012) can come out as an alternative for the standardisation of taxonomies and procedures to annotate speech act corpora. In (Fang et al., 2012), in order to deal with these negative points, SwDA was annotated with the ISO standard and an evaluation of the taxonomy applicability was made. In (Mezza et al., 2018), several benchmark schemes are mapped to ISO, with the same purpose as the previous work. These efforts, in turn, illustrate the interest of current research in producing tagsets of speech acts that can be compared across corpora.

Beyond taxonomic diversity, another characteristic presented by most available corpora is that they usually do not suffer from high class imbalance, there being a few, if any, examples of databases where some speech act class lies highly predominant in relation to others. This, however, is characteristic to the the journalistic field, in which the speech act 'inform' corresponds to over 90% of the examples in our corpus (see Section 3.2). In SwDA and MRDA, for example, the majority class corresponds to about 36% and 60% of the examples, respectively.

3 Materials and Methods

In this work, we build on the Porttinari-base corpus, in its August 10, 2022 version. Comprising 8,420 sentences (168,399 tokens) from 1,073 news texts written in (Brazilian) Portuguese, the corpus was mannualy annotated with (morpho)syntactic features, under the Universal Dependencies (UD) (Nivre et al., 2020) paradigm. Its construction followed a five-stage pipeline, comprising Plain Text Preparation, Automatic PoS (Part of Speech) Tag Annotation, Manual PoS Tags Revision, Semi Automatic Lemma Annotation and Semi Automatic PoS Annotation, as described in detail in (Lopes et al., 2022).

This corpus was selected for this research because it is exclusively in Brazilian Portuguese, following the standard norm of the language, also being annotated with UD PoS tags. An annotation example in Porttinari-base can be seen in Table 1. In this table, it is possible to notice the news text segmented into sentences along with the PoS tags assigned to each token in the sentence.

For the manual annotation of the Porttinari-base corpus with speech acts, a random sample of 50% of its news was selected, totalling 536 news (4,091 sentences). Data selection was based on the news and not on individual sentences due to the importance of context for the task. All sentences from the selected news texts were then annotated by one of the researchers. In doing so, our intent was to preserve the remaining 50% so that an automatic classifier could be trained in the annotated half and applied to the rest of the corpus in a transductive manner.

3.1 Speech Acts Taxonomy

As mentioned, several different Speech Act tagsets are currently used by extant research. This leads to some negative aspects, such as the difficulty in comparing different studies, the lack of standardisation of label meaning (such as the use of the same label with different meanings or different labels for the same speech act), the use of very specific tagsets, highly tailored to certain tasks, which makes their reuse difficult, the lack of consensus on a hierarchy of speech acts, and the existence of Speech Acts that are not reusable across different tasks.

Eu(PRON) sei(VERB) que(SCONJ) tô(AUX) lascado(ADJ) ,(PUNCT) todo(DET) dia(NOUN) tem(VERB) um(DET) processo(NOUN) .(PUNCT) *('I know I'm screwed up, every day there's a lawsuit.')*
Eu(PRON) não(ADV) quero(VERB) nem(ADV) que(SCONJ) Moro(PROPN) me(PRON) absolva(VERB) ,(PUNCT) eu(PRON) só(ADV) quero(VERB) que(SCONJ) ele(PRON) peça(VERB) desculpas(NOUN) ,(PUNCT) disse(VERB) Lula(PROPN) durante(ADP) um(DET) seminário(NOUN) sobre(ADP) educação(NOUN) em(ADP) Brasília(PROPN) .(PUNCT) *(I don't even want Moro to absolve me, I just want him to apologize, said Lula during a seminar about education in Brasilia.)*

Table 1: Examples of Porttinari-base sentences and their corresponding PoS tags.

In an attempt to solve this problem, ISO 24617-2 (ISO, 2012) was proposed as a standard for annotating Speech Acts in different domains. For this reason, in this work we decided to use this tagset of Speech Acts. ISO 24617-2's taxonomy is composed of 56 communicative functions, divided in 9 dimensions (Allo and Auto Feedback, Turn Managment, Time Managment, Discourse Structuring, Own and Partner Communication Managment, Social Obligation Management and Contact Management). Dimensions are classes of Dialogue Acts referring to a particular category of semantic content (type of information, situation, action, event or objects that form the semantic content of a dialogue act), according to a particular aspect of communication.

In addition to dimensions, communicative functions are also divided in two groups – General Purpose and Specific Purpose. The General Purpose group refers to functions that can be used with any type of semantic content, with the main characteristic of obtaining or requesting information and discussing actions. On the other hand, Specific Purpose functions deal only with the category of semantic content related to their dimension, encompassing Speech Acts that are divided according to their specific dimensions.

Figure 1 lists the dimensions and communicative functions defined by ISO 24617-2, separated according to their type and dimension. In bold, we highlight the communicative functions that are more in line with the journalistic nature of our corpus. Table 2 presents some examples of sentences from the Porttinari-base corpus and their respective communicative functions. For more examples, we refer the interested reader to da Silva et al. (2023). As expected, many communicative functions are more tailored to dialogues, being of limited use to other genres.

3.2 Corpus Annotation

The annotation procedure followed two steps: (i) dimension identification and (ii) communicative function identification, based on the ISO 24617-2 descriptions and our considerations and adaptations to the journalistic nature of Porttinari-base. With that in mind, the most appropriate and sentence-specific communicative function was selected for each sentence. Each sentence in the sample was necessarily annotated with one speech act. For a detailed description of the annotation procedure we refer the interested reader to da Silva et al. (2023).

As an example, consider the sentence *"Isso é uma vergonha para os nova-iorquinos."* (*"That's a shame for New Yorkers."*). At first glance, its communicative function might be "inform". It might, however, also be a "disagreement". In such cases, we always assign the most specific communicative function (in this case, "disagreement") to the sentence.

There are also cases where more than one label fits the sentence, as in *"Até resolver esse problema filosófico, convém continuar a investir em métodos anticoncepcionais."* (*"Until this philosophical problem is solved, it is advisable to keep on investing in contraceptive methods."*). In this case, both communicative functions "inform" and "suggestion" are adequate, and we leave to the annotator decide which label to adopt.

Finally, another point of attention is the attribution of communicative functions to sentences that describe or inform about some speech act. For example, the sentence *"Ao saber que teria que abandonar a prova, Vettel pediu desculpas à equipe."* (*"Upon knowing that he would have to leave the race, Vettel apologised to the team."*) describes an apology. However, it does not have the communicative function of an apology, merely informing about this act instead.

General purpose communicative functions

Information providing functions	**Information seeking functions**	**Comissive functions**	**Directive functions**
1. inform **2. agreement** **3. disagreement** **4. confirm** **5. disconfirm** **6. correction** **7. answer**	1. **question** 2. propositonal question 3. set question 4. check question 5. choice question 6. test question	1. **promise** 2. offer 3. address request 4. accept request 5. decline request 6. address suggest 7. accept suggest 8. decline suggest	1. **instruct** 2. **request** 3. **suggest** 4. address offer 5. accept offer 6. decline offer

Dimension Specific communicative functions

Feedback functions	Turn-management functions	Time-management functions	**Social obligations management functions**
1. auto positive 2. auto negative 3. allo positive 4. allo negative 5. feedback elicitation	1. turn accept 2. turn take 3. turn grab 4. turn assign 5. turn release 6. turn keep	1. stalling 2. pausing	1. **apology** 2. **thanking** 3. **compliment** 4. **congratulation** 5. **sympathy expression** 6. init greeting 7. return greeting 8. init self introduction 9. return self introduction 10. accept apology 11. accept thaking 12. init goodbye 13. return goodbye
Discourse-structuring functions	Own and partner management functions	Contact management functions	
1. interaction structuring 2. opening 3. topic shift	1. self error 2. retraction 3. self correction 4. completion 5. correct misspeaking	1. contact check 2. contact indication	

Figure 1: General purpose and dimension specific communicative functions defined by ISO 24617-2

Function	Sentence
inform	Tite says he wants to remain in the national team after the World Cup in Russia. Tite diz querer seguir em a seleção após o Mundial de a Rússia.
question	Where does this icon go in the future? Para onde esse ícone vai em o futuro?
suggest	Wash your car in the shade so that the chemicals don't cause stains. Lave o carro na sombra, para que as substâncias químicas não causem manchas.
disagreement	This thing about the best ice cream in the world is nonsense. Esse negócio de melhor sorvete de o mundo é bobagem.
disconfirm	What is said is not true, that I had no right to leave the country. Não é verdade o que se fala, que eu não tinha o direito de sair de o país.

Table 2: Examples of some communicative functions.

In the end, the selected sample from Porttinari-base (with its 536 news and 4,091 sentences), was manually annotated with speech acts according to the communicative functions proposed by ISO 24617-2. To illustrate, Table 3 presents the sentences from Table 1, with their corresponding speech acts.

Speech acts distribution across the corpus can be seen in Table 4. In this table, we observe the great imbalance of speech act classes in this corpus, with a clear prevalence of the class 'inform'. This comes as no surprise, given the journalistic nature of the corpus. Moreover, it can be noted that many communicative functions defined by ISO's taxonomy could not be identified during the annotation procedure. This is due to the fact that the speech acts defined by ISO were formulated mainly for dialogues, which makes labels strongly related to dialogues of little use when it comes to news.

This imbalance becomes even more prominent as we climb up the taxonomy's hierarchy, as shown in Table 5. In this table, categories were grouped by type and dimension of their communicative functions. As expected, most of the functions are of general use, with 'Social Obligations' figuring as the only dimension of the Specific type. This, in turn, was found mainly in news containing interviews in the form of dialogues.

3.3 Experimental Design

In this experiment, we fine tuned BERTimbau (Souza et al., 2020), in the annotated sample corpus described above, to the task of speech act classification. We then randomly split 64% of the data set for training, with 20% being held for testing and 16% for validation purposes. Since we opted for stratified sampling, some of the classes[3] could not be included in all sets. These were then removed, which resulted in a total of 13 classes being used during this procedure.

The experiment was run in Google Colab, with 12GB of RAM, 100GB of disk space and a Tesla T4 GPU with 15GB of RAM. It was performed using the Pytorch library[4] and the large version of BERTimbau[5]. We used a training batch size with 32 examples, as this is the largest size supported by the hosting machine, varying the training epochs from 1 to 5.

To deal with class imbalance, we performed experiments adding different weights to both the majority and minority classes in the model's cross entropy loss function to generate a higher penalty for model errors in minority classes. The weight defined for each class was inversely proportional to their respective frequencies in the validation set. Competing models were then evaluated in the validation set, and the best combination of hyper-parameters (the number of epochs and the existence or not of weights in the loss function) was used to build the final model, which was then retrained in the combination of the training and validation sets (which comprised 80% of the annotated corpus) and finally assessed in the test set. This code implementation is publicly available [6] to the community.

4 Results and Discussion

Figures 2 to 4 present accuracy, weighted averaged F1[7] and macro averaged F1[8], respectively, as a result from the fine tuning of BERTimbau along all epochs, measured in the validation set. The figures also distinguish between the application or not of weights in the cost function (*use_weight* in the figures). As expected, both accuracy and weighted F1 present much higher values than macro F1, given the severe imbalance of classes in the corpus.

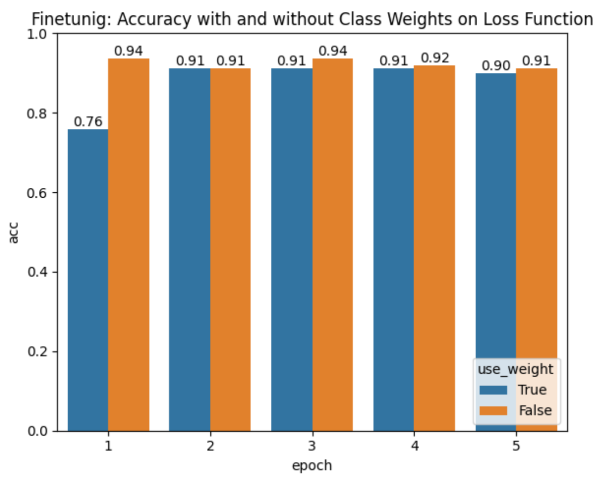

Figure 2: Accuracy in each epoch at the validation set.

In this case, a maximum of 92% weighted F1 could be observed at the third epoch, with no

[3] correction, agreement, congratulation and apology classes were not used in the experiment due to their lack of examples

[4] https://pytorch.org/

[5] https://huggingface.co/neuralmind/bert-large-portuguese-cased

[6] https://github.com/natalypatti/porttinari-base-speech-acts

[7] Average of the F1 obtained in each class, weighted by the proportion of classes in the set.

[8] Arithmetic mean of per class F1.

Sentence	Type	Dimension	Function
‘I know I’m screwed up, every day there’s a lawsuit.’	General	information providing	disagreement
I don’t even want Moro to absolve me, I just want him to apologize, said Lula during a seminar about education in Brasilia.	General	information providing	inform

Table 3: Porttinari-base annotation examples

Class	Total	(%)	Class	Total	(%)
inform	3725	91.054	sympathy Exp	11	0.269
question	96	2.347	request	7	0.171
suggest	64	1.564	confirm	6	0.147
disagreement	62	1.516	promise	6	0.147
disconfirm	26	0.636	correction	4	0.098
compliment	24	0.587	agreement	4	0.098
answer	22	0.538	congratulation	2	0.049
instruct	18	0.440	apology	1	0.024
thanking	13	0.318			

Table 4: Annotated Speech Act classes in the Porttinari-base corpus

Type	Dimension	Count	(%)
General	information-providing functions	3849	94.08
	information-seeking functions	97	2.37
	directive functions	88	2.15
	commissive functions	6	0.14
Specific	social obligations functions	51	1.24

Table 5: Types and dimensions of Speech Acts in the Porttinari-base corpus

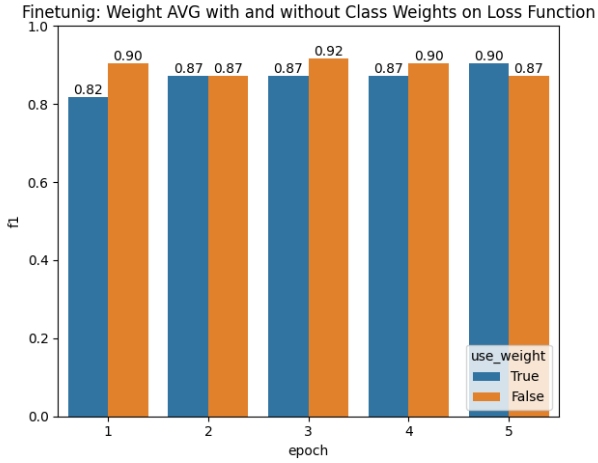

Figure 3: Weighted F1 in each epoch at the validation set.

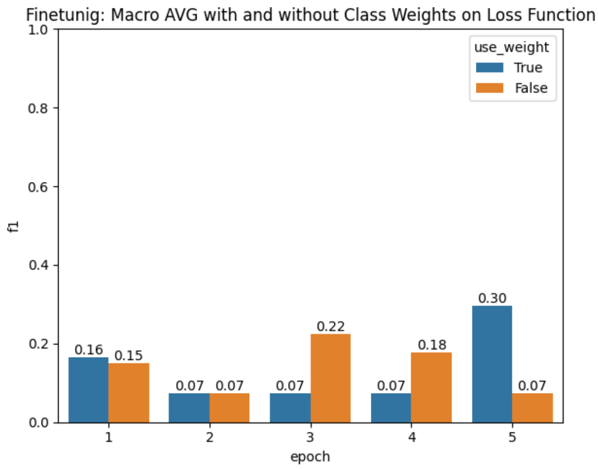

Figure 4: Macro F1 in each epoch at the validation set.

weights applied to the loss function. Accuracy also topped at the third epoch without weights, at 94%. These high values reflect the good performance of the model mainly in predicting the majority class (*i.e.* 'inform').

When it comes to macro averaged F1, however, figures drop substantially, since the model's performance in the remaining classes becomes more evident. In this case, the best values are found in the fifth epoch, for the weighted version (30%) and once again at the third epoch (22%), in the unweighted version of the loss function.

The hyperparameters from the best macro averaged F1 (*i.e.* from the fifth epoch with weights) were then used in the final model, which was once again fine tuned, but this time in a combination of training and validation sets. The results of testing this final model at the test set can be seen in Table 6. As it turns out, although smaller, macro averaged F1 did not differ so much when compared to the validation set.

Accuracy (%)	91.6
Weighted Averaged F1 (%)	91.4
Macro Averaged F1 (%)	29.5
Examples	816

Table 6: Results at the test set, with 5 epochs and weights in the cost function

A breakdown of the model's performance across classes can be seen in Table 7, which confirms its higher performance at the majority class (in this case, 'inform', with a 95.7% F1). Interestingly, despite the low number of examples (19), 'question' also delivered a high F1 (92.6%), which could be an indication of how easily it can be recognised by this model. At the other end of the scale, its performance dropped to nil when dealing with classes with but a few examples in the corpus (tipically, less than 4), with the exception of 'instruct' which, despite having only four examples in the corpus, could deliver a 44% F1.

As expected, class imbalance posed a great challenge for the model in terms of F1. Still, the existence of outlying classes, such as 'compliment', 'disconfirm' and 'instruct' which, despite being rare, can still be recognised by the model, calls for a deeper linguistic analysis as to why this was the case.

4.1 Limitations to this Work

Considering the results described in Section 3.3, we observe that the fine tuning of BERTimbau, even with the help of class weights in the cost function, was not sufficient to satisfactorily address the classification of minority speech act classes. Future research directions could be to employ more features in classification such as, for example, context and syntactic features (Liu et al., 2017; Blache et al., 2020), which might help with this issue.

Another important drawback of this research lies in the fact that the annotation process was carried out by one annotator only, which can generate a bias towards this annotator's personal opinion, thereby limiting the generalisation of the resulting classification. Although we believe this limitation not to decrease the value of the resource as a whole, we intend to deal with it in a follow-up version of the corpus.

Class	Precision (%)	Recall (%)	F1 (%)	Examples
inform	96.0	95.4	95.7	745
question	86.3	100	92.6	19
suggest	34.7	61.5	44.4	13
disagreement	33.3	41.6	37.0	12
compliment	66.6	40.0	50.0	5
disconfirm	20.0	20.0	20.0	5
answer	0	0	0	4
instruct	40.0	50.0	44.0	4
thanking	0	0	0	3
request	0	0	0	2
sympathyExpress	0	0	0	2
confirm	0	0	0	1
promise	0	0	0	1

Table 7: Detailed performance of BERTimbau Finetuning using 5 epochs

5 Conclusion

In this work, we presented an annotated subset of the Porttinari-base corpus, manually labeled with Speech Acts according to the taxonomy proposed by ISO 24617-2. This is the first corpus, to the best of our knowledge, in Brazilian Portuguese annotated with Speech Acts, being also probably the first in the journalistic field.

With this corpus, we were able to verify the challenge related to dealing with the automatic identification of speech acts in news texts, given their high class imbalance, where "inform" dominates the scenario with over 90% of the sentences. In the experiment carried out by fine tuning BERTimbau, we noticed the good model performance in the classification of the predominant class and its difficulty in the less frequent classes. Despite this difficulty, the model still managed to get hits in these more challenging classes, encouraging new efforts to delve deeper into this issue.

We hope that this corpus, which is freely available[9] to the community under a Creative Commons license, may contribute to the field, especially to research focused on Brazilian Portuguese. As for directions for future work, we intend to proceed with the complete annotation of Porttinari-base, by applying an automatic transductive learning algorithm, taking our current annotated corpus as a start point. Another interesting venue for future research would be to try to add some syntactic information to the model, since this could be useful for differentiating some classes of speech acts (*cf.* Liu et al., 2017; Blache et al., 2020).

[9] https://github.com/natalypatti/porttinari-base-speech-acts

Acknowledgments

This work was carried out at the Center for Artificial Intelligence of the University of São Paulo (C4AI - http://c4ai.inova.usp.br/), with support by the São Paulo Research Foundation (FAPESP grant #2019/07665-4) and by the IBM Corporation. The project was also supported by the Ministry of Science, Technology and Innovation, with resources of Law N. 8.248, of October 23, 1991, within the scope of PPI-SOFTEX, coordinated by Softex and published as Residence in TIC 13, DOU 01245.010222/2022-44.

References

John Langshaw Austin. 1962. *How to do things with words*, 1 edition. Oxford University Press.

Emily M. Bender and Alexander Koller. 2020. Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5185–5198, Online. Association for Computational Linguistics.

Philippe Blache, Massina Abderrahmane, Stéphane Rauzy, Magalie Ochs, and Houda Oufaida. 2020. Two-level classification for dialogue act recognition in task-oriented dialogues. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 4915–4925, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Lin Chen and Barbara Di Eugenio. 2013. Multimodality and dialogue act classification in the RoboHelper project. In *Proceedings of the SIGDIAL 2013 Conference*, pages 183–192, Metz, France. Association for Computational Linguistics.

Mark Core and James Allen. 2001. Coding dialogs with the damsl annotation scheme.

Nataly Leopoldina Patti da Silva, Norton Trevisan Roman, and Ariani Di Felippo. 2023. Manual de anotação do corpus porttinari-base com atos de fala. Technical report, University of São Paulo.

A. Fang, J. Cao, H.C. Bunt, and X. Liu. 2012. The annotation of the switchboard corpus with the new iso standard for dialogue act analysis. In *Proceedings of the 8th joint ISO-ACL Sigsem workshop on interoperable semantic annotation, Pisa*, pages 13–18. ILC-CNR.

J.J. Godfrey, E.C. Holliman, and J. McDaniel. 1992. Switchboard: telephone speech corpus for research and development. In *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 517–520 vol.1.

ISO. 2012. Iso 24617-2:2012: Language resource management – semantic annotation framework (semaf) – part 2: Dialogue acts.

Dan Jurafsky and Elizabeth Shriberg. 1997. Switchboard swbd-damsl shallow-discourse-function annotation coders manual.

Yang Liu, Kun Han, Zhao Tan, and Yun Lei. 2017. Using context information for dialog act classification in DNN framework. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2170–2178, Copenhagen, Denmark. Association for Computational Linguistics.

Lucelene Lopes, Magali Sanches Duran, Maria das Graças Volpe Nunes, and Thiago Alexandre Salgueiro Pardo. 2022. Corpora building process according to the universal dependencies model: an experiment for portuguese.

Stefano Mezza, Alessandra Cervone, Evgeny Stepanov, Giuliano Tortoreto, and Giuseppe Riccardi. 2018. ISO-standard domain-independent dialogue act tagging for conversational agents. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3539–3551, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Jan Hajič, Christopher D. Manning, Sampo Pyysalo, Sebastian Schuster, Francis Tyers, and Daniel Zeman. 2020. Universal Dependencies v2: An evergrowing multilingual treebank collection. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 4034–4043, Marseille, France. European Language Resources Association.

Elizabeth Shriberg, Raj Dhillon, Sonali Bhagat, Jeremy Ang, and Hannah Carvey. 2004. The ICSI meeting recorder dialog act (MRDA) corpus. In *Proceedings of the 5th SIGdial Workshop on Discourse and Dialogue at HLT-NAACL 2004*, pages 97–100, Cambridge, Massachusetts, USA. Association for Computational Linguistics.

Fábio Souza, Rodrigo Nogueira, and Roberto Lotufo. 2020. Bertimbau: Pretrained bert models for brazilian portuguese. In *Intelligent Systems*, pages 403–417, Cham. Springer International Publishing.

Henry S. Thompson, Anne Anderson, Ellen Gurman Bard, Gwyneth Doherty-Sneddon, Alison Newlands, and Cathy Sotillo. 1993. The HCRC map task corpus: Natural dialogue for speech recognition. In *Human Language Technology: Proceedings of a Workshop Held at Plainsboro, New Jersey, March 21-24, 1993*.