Survey: Reduplication vs Repetition classification

Arif Ahmad, Preethi Jyothi and **Pushpak Bhattacharyya**
Centre for Machine Intelligence and Data Science (C-MInDS)
Indian Institute of Technology Bombay
{arifahmad, pjyothi, pb}@cse.iitb.ac.in

Abstract

Reduplication is a common linguistic phenomenon in many languages, including Indian languages, where it serves a distinct grammatical and semantic function. However, in the context of automatic speech recognition (ASR) transcripts, reduplication can often be mistaken for repetition, leading to errors in downstream NLP systems that rely on these transcripts as input. The unavailability of labeled data for this task further impairs progress. In this survey paper, we propose the task to classify reduplication vs. repetition in ASR transcripts to improve the performance of disfluency correction systems and downstream NLP tasks. We also clearly deine the problem in hand and the motivation behind the task. We discuss about existing ASR datasets that can be used to create new datasets for the task of reduplication vs repetition classification.

1 Introduction

Reduplication is a linguistic phenomenon that involves the repetition of all or part of a word to create a new word with a different meaning or to intensify the original word's meaning. This phenomenon is prevalent in many Indian languages and is used to create plurals, express tense or aspect, or create new words with different shades of meaning. However, reduplication can often be confused with repetition, a type of disfluency in spontaneous speech. Repetitions can significantly inhibit the performance of downstream NLP tasks, including machine translation. Disfluency detection and correction are frequently employed as preprocessing steps to identify and remove such disfluent words.

While disfluency detection has been studied extensively in English, it has received significantly less attention in other languages, including Indian languages. This lack of attention is especially concerning as India is a land of many languages, with over 700+ languages spoken nationwide. Providing healthcare, legal, and tourist information in various Indian languages is essential to ensure access to critical information for the population. Machine Translation provides a promising solution to the challenges of manual translation, but the availability of Machine Translation technology is limited. Particularly in rural areas where illiteracy rates are higher than in the rest of India, a more accessible mode of access to Translation technology is required. Speech-to-Speech Machine Translation systems provide a more accessible mode of interaction, but they struggle with disfluencies often present in spontaneous speech.

Therefore, this study focuses on discriminating reduplication from repetition in Indian languages. We aim to develop a model that can effectively classify reduplication and repetition to improve the performance of Disfluency correction systems and, in turn, downstream NLP systems that use ASR transcripts as input. Specifically, we investigate the effectiveness of different features and models for the task of reduplication vs. repetition classification in Indian languages. We also analyze the impact of different types of reduplication and repetition on classification performance. Our study will provide insights into the nature of reduplication and repetition in Indian languages and inform the development of more effective Disfluency correction systems and Speech-to-Speech Machine Translation systems.

2 Problem Statement

In this study, we address the problem of discriminating between reduplication and repetition in Indian languages given a speech utterance transcript. Reduplication is a linguistic phenomenon where a part or the entirety of a word is repeated to create a new word or intensify the meaning of the original word. It can occur at the beginning, middle, or end of a word, and is a commonly used phenomenon in

many Indian languages. On the other hand, repetition involves the exact repetition of a word without any additional meaning. The distinction between reduplication and repetition is important for many downstream NLP tasks such as speech-to-speech machine translation, disfluency detection, and speech synthesis. While reduplication and repetition have been studied in English and other languages seperately, very little research has been done in the context of Indian languages. We aim to develop models that can accurately distinguish between reduplication and repetition in Indian languages, which can benefit a wide range of NLP applications.

3 Motivation

Reduplication and repetition are common phenomena observed in Indian languages. They involve repeating a word or a portion of it to create new forms, often with a different meaning or connotation. For instance, in Hindi, the word "paani" (water) can be reduplicated to form "pani-pani," which means "a lot of water." Similarly, in Telugu, the word "koti" (crore) can be repeated to form "koti-koti," which means "innumerable." Reduplication and repetition are extensively used in daily conversations, poetry, and literature in Indian languages.

However, despite their widespread usage, reduplication and repetition are not well-studied together in computational linguistics. There is a lack of research on how these phenomena can be leveraged to improve natural language processing tasks in Indian languages. Additionally, the distinction between reduplication and repetition is not always clear, and there is a need for a systematic investigation of their similarities and differences.

This motivates us to study reduplication and repetition in Indian languages from a computational linguistics perspective. Our goal is to explore their patterns, meanings, and usage in different contexts and to investigate how they can be effectively utilized in natural language processing tasks such as sentiment analysis, machine translation, and text generation. Our research will contribute to a better understanding of the unique features of Indian languages and pave the way for the development of more effective natural language processing systems for these languages.

4 Disfluency

Speech utterances are classified into two types: read speech and conversational speech. The term “read speech" refers to utterances in which the speaker reads the information to be said from a source. Conversational speech is natural and spontaneous, with the speaker thinking and articulating as they speak. Spontaneous or conversational speech, may contain irregularities. One of the irregularities is *disfluency*. Disfluencies are words that appear in conversational speech but offer no semantic sense to the phrase. Speakers often use filler words, repeat fluent phrases, suddenly change the content of speech, and make corrections to their statements. These are some common disfluencies. For example, below are a few disfluent sentences with the disfluencies highlighted in italics and bold:

Show flights from Boston on ***uh*** from Denver on Monday.

Let us, I mean, let me work on the problem.

Well, this is this is ***you know*** a good plan.

Disflencies have often been viewed as noisy and irregular events. While people generally don’t even notice disfluencies in day-to-day conversation, early foundational work in computational linguistics demonstrated how common they are. In 1994, using the Switchboard corpus, (Shriberg, 1994) shows that they appear remarkably regularly in conversational speech. In fact, (Shriberg, 1994) shows that for a sentence of 10-13 words, there is a 50% probability that it contains a disfluency. This probability increases as the length of the sentence increases. Figure 1 shows a plot of the proabibilty that a sentence is disfluent vs. the length of the sentence.

When disfluent words are present in a sentence, they can add complexity, reduce semantic clarity, and render the phrase non-fluent; which can be detremental to downstream applications (such as translation) that use the captured disfluent speech.

4.1 Structure of Disfluency

Multiple studies in conversation analysis, psycholinguistics and computational linguistics such as ((Hockett, 1973), (Schegloff, 1987), (Goodwin, 1986), (Levelt, 1983), (Hindle, 1983)) have independently found the same surface form in a majority of disfluencies. The disfluencies start with the material that will ultimately be replaced, followed

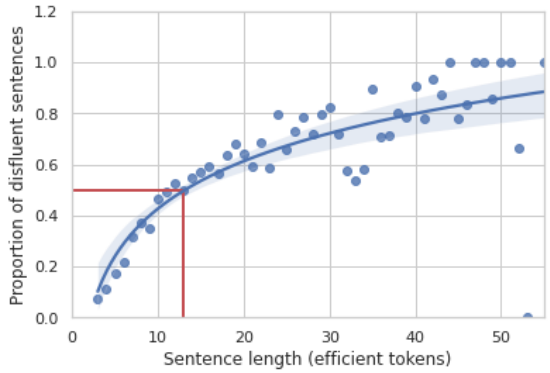

Figure 1: The proportion of sentences from the Switchboard dataset with at least one disfluency plotted against sentence length measured in non-disfluent (i.e., efficient) tokens in the sentence. (Shriberg, 1994)

by optionally one or more editing phrases (such as well" or "I'm sorry), ending the replacing material. These regions are continuous, and removing a continuous length of material containing the error yields in the expected intended utterance.

(Shriberg, 1994) shows that a disfluency consists of three components: reparandum (followed by interruption point), then interregnum, followed by repair. Figure 2 shows an example which contains all the above mentioned components of disfluencies.

Figure 2: Surface Structure of Disfluency

It is necessary to keep in mind that none of the above parts are required to be present in every disfluency. But a disfluent phrase, must include at least one component. Below are the details of what each component indicates in a disfluency:

- **Reparandum** contains words that were not initially intended to be in the speech. As a result, this section contains one or more words that will be repeated or corrected (in the event of Repetition or Correction) or dropped entirely (in case of a False Start).
- The **interruption point** follows, signalling the conclusion of the reparandum. It has nothing to do with any sort of pause or auditory event.
- The **interregnum** comes after it. This section consists of an editing word, a non-lexicalized filler pause such as "uh," "um," or discourse markers such as "well," "you know," or interjections, or just an empty pause, i.e., a brief period of stillness.
- The final stage is **repair**. In the repair step, words from the reparandum are eventually repaired or repeated (in the case of Repetition or Correction), or an entirely new phrase is started (in the case of False Start).

Figure 3: Disfluencies with empty interregnum

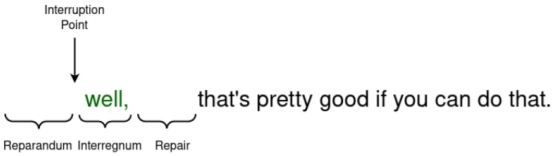

Figure 4: Disfluencies with only interregnum

Figures 3 and 5 show examples where on of the components of the disfluency is missing. In the example in Figure 3, the reparandum is immediately followed by the repair and the interregnum does not show up as a textual cue in the sentence. Similarly, in the example in Figure 5, the reparandum and repair are not present, but the speaker talks disfluently due to the presence of interregnum.

4.2 Types of Disfluency

There are various types of disfluency phenomena which are not captured in the transcripts of the text captured by ASR systems, such as unfilled pauses and uncorrected prosodic errors. These are beyond the scope of the present work. We hence look into only those types of disfluencies in which a contiguous stretch of linguistic material must be deleted to arrive at the sequence the speaker "intended, likely the one that would be uttered upon a request for repetition. Such cases include a wide class of phenomena commonly referred to as filled pauses, rep-

etitions, false starts, repairs, and a variety of other terms.

According to (Honal and Schultz, 2003), disfluencies are classified into two types: simple and complex disfluencies. Simple disfluencies include filled pauses like “uh," “ah," and “um," as well as discourse markers like “yeah," “well," “you know," and “okay." This category also includes interjections like “oops" and “ugh." Many times, words like yeah, okay are also marked as filled pauses.

In repetition or correction, the abandoned phrase is repeated with just minor or no modifications in syntactical structure. On the other hand, it is a false start if an entirely other syntactical structure with a distinct meaning is started following the abandoned sentence. Edit occurs to convey that the words said earlier were not meant. The table below depicts the many sorts of disfluencies and an example of each in English -

5 Reduplication

The act of repeating all or part of a word for emphasis or to communicate a meaning is known as reduplication. It is common in Indian languages; some examples of reduplication (Montaut, 2009) in Hindi are given below:

देश में **जगह जगह** पर पुरानी इमारतें है |
यह लो तुम्हारी चाय. **गरम गरम** है, पियो |
सच सच बताओ |

In the context of disfluencies, we should highlight that reduplication might be confused with the repetition disfluency type. Reduplications are grammatically valid deliberate repeats that should not be classified as disfluencies. Hence we discuss about reduplications in detail here.

Reduplication, unlike repetition, is a morphological process where the whole word or part of it (root or stem) is repeated exactly or with some change. Its importance is best summarized in (Sapir, 2004) as, “Nothing is more natural (in language) than the prevalence of reduplication, in other words, the repetition of all or part of the radical element". Semantically, reduplication is used to indicate concepts such as distribution, plurality, frequency, customary activity, increase of size, added intensity, continuance (Sapir, 2004). However, it should be noted that not all of these concepts are expressed as reduplication in a given language, and their usage can vary from language to language. Reduplication is found in a wide variety of languages, but here we discuss about reduplication only in English and Hindi languages.

5.1 Reduplication in English

There are several types of reduplication in English, ranging from casual expressive vocabulary (rhyming, exact, ablaut, shm) to grammatically meaningful forms (comparitive and contrastive focus). These are decribed below:

- **Rhyming reduplication:** When the repeated word rhymes with the original word, the repetition is called a rhyming compound, or a rhyming reduplicative.

  Examples: Easy-peasy, hocus-pocus, hokey-pokey, itsy-bitsy, super-duper, etc.

- **Exact reduplication:** When the repeated word rhymes with the original word, it is called an exact reduplication. Exact reduplications can be employed to emphasise thenintensity of a word in several versions of English ("He wants it now now"); in South African English, ’now-now’ means’relatively soon.’

  Examples: Bye-bye, Aye-aye, back-to-back, blah-blah, boo-boo, night-night, pom-pom.

- **Ablaut reduplication:** In ablaut reduplications, the first vowel is almost always a high vowel (typically as in hit) and the reduplicated vowel is a low vowel (typically æ as in cat or as in top).

  Examples: Chit-chat, flip-flop, hip-hop, jibber-jabber, ping-pong, sing-song, tick-tock

- **Shm-reduplication:** This type of reduplication can be used with almost any word. Shm-reduplication is more commonly found in informal conversations and not formal text or communication.

  Examples: baby-shmaby, cancer-shmancer and fancy-shmancy.

- **Comparative reduplication:** The reduplication of the comparative in the sentence "The meadows got greener and greener" suggests that the comparative is getting increasingly true over time, essentially meaning "The meadows appeared progressively greener as

Type	Description	Constituents	Examples
Filled Pause	Non lexicalized sounds with no semantic content.	uh, um, ah, etc	but **uh** we have to go through the same thing.
Interjection	A restricted group of non lexicalized sounds indicating affirmation or negation.	uh-huh, mhm, uh-uh, ugh, uh-oh, oops etc.	**Oops,** I did not know that you would get hurt.
Discourse Marker	Words that are related to the structure of the discourse in so far that they help beginning or keeping a turn or serve as acknowledgment. They do not contribute to the semantic content of the discourse.	well, you know, okay, yeah etc.	**Well,** this is a good plan.
Repetition or Correction	Exact repetition or correction of words previously uttered. A correction may involve substitutions, deletions or insertions of words. However, the correction continues with the same idea or train of thought started previously.		If I **can't** don't know the answer myself, I will find it.
False Start	An utterance is aborted and restarted with a new idea or train of thought.		**We'll never find a day** what about next month?
Edit	Phrases of words which occur after that part of a disfluency which is repeated or corrected afterwards or even abandoned completely. They refer explicitly to the words which just previously have been said indicating that they are not intended to belong to the utterance.		We need two tickets, **I'm sorry**, three tickets for the flight to Boston.

Figure 5: Types of Disfluencies with description and example (Honal and Schultz, 2003)

time went on." The comparison in reduplication is of the thing being compared to itself throughout time. This construction is common in speech, even in formal situations, although it is less prevalent in formal written texts.

- **Contrastive focus reduplication:** Exact reduplication can be used in conjunction with contrastive emphasis (usually when the first word is stressed) to imply a literal, rather than metaphorical, example of a noun, or possibly a type of Platonic ideal of the noun, as in "Is it carrot cheesecake or carrot CAKE-cake?" This is comparable to the above described Finnish use. It is also used to contrast "genuine" or "pure" goods with imitations or less pure versions. In a coffee establishment, for example, one may be asked, "Do you want soy milk?" and answer, "No, I want milk milk." This suggests that they desire "genuine" milk.

5.2 Reduplication in Hindi

In Hindi, certain parts of speech, such as nouns, adjectives, adverbs, verb stems, interrogatives, etc., may be reduplicated (Jain, 2007) to indicate the following:

- intensify the meaning of word or phrase

  Example: **जल्दी जल्दी** काम खत्म करो (Finish your work **quickly**)

- distribute an item, attribute, or quality throughout a group of **specific people** across time and space

  Example:

  फ्रांस के **कोने कोने** में बेकरी हैं (There are bakeries in **every corner** of France)

  . . . Distributed across space

  यह गीता **सालों सालों** से हमारे परिवार में रही है (This Geeta has been in our family **for years**.)

  . . . Distributed across time

  कोरिया में **बड़ी बड़ी** आँखों को सुन्दर माना जाता है (**Big** eyes are considered beautiful in Korea.)

  . . . Distributed quality among people/objects

- to express frequency or thoroughness of action

  Example:

  मार्वल मूवीज के बारे में **सुन सुनकर** उससे रहा नहीं गया और वह भी फिल्म देखने चला गया (After **repeatedly hearing** about Marvel movies, he could not stay and he also went to see the movie) . . . frequency

- ask for details (especially in interrogative sentences)

  Example:

  यूरोप में आप **कहाँ कहाँ** घूमे ? (**Where all** did you travel in Europe?)

6 Datasets

To perform our task of distinguishing reduplication and repetition, we require speech corpora that are annotated with disfluencies along with the transcriptions of the speech. However, as discussed, due to disfluencies often being treated as noisy and irregular events, we often donot find annotations of disfluencies in all speech corpora. Here, we mention some available speech corpora with disfluencies labelled along with synthetic methods to add disfluencies to datasets without disfluenices.

6.1 Available datasets with disfluencies

Some of the common datasets which contain disfluencies and have traditionally been used for disfluency detection and correction are:

- SWITCHBOARD (SWBD) corpus: This is a corpus of informal human-human telephone conversations on various topics (Godfrey et al., 1992)

- AMERICAN EXPRESS/SRI corpus (also known as AMEX corpus): A corpus of human-human air travel planning dialogs (Kowtko and Price, 1989)

- ATIS corpus: human-computer dialog in the air travel planning domain (Dahl et al., 1994)

More recently, Disfl-QA (Gupta et al., 2021a) is a dataset released for disfluency correction. Disfl-QA is a derivative of SQUAD ((Rajpurkar et al., 2016), (Rajpurkar et al., 2018)) with textual disfluencies in previously fluent questions .

6.1.1 SWITCHBOARD (SWBD) corpus

The SWBD corpus comprises almost three million words from over 2430 telephone calls on diverse themes. The corpus was compiled at Texas Instruments and is made available by the LDC. More details can be found in (Godfrey et al., 1992), we discuss here specifically about the disfluent data in the SWBD corpus. The dataset consists of 1694 disfluencies of which 1574 are basic disfluencies while 120 are complex disluencies. There are 1227 disfluent sentences in the corpus.

6.1.2 AMEX corpus

Telephone conversations between SRI workers and American Express travel agents comprise the AMEX corpus of human-human air travel planning dialogues. There was no task assigned in this corpus; rather, telephone conversations between SRI employees and American Express travel agents (i.e., calls containing real trip plans) were tape-recorded once agents acquired the employee's permission at the start of the call. More details can be found in (Kowtko and Price, 1989), we discuss here specifically about the disfluent data in the AMEX corpus. The dataset consists of

745 disfluencies of which 672 are basic disfluencies while 73 are complex disluencies. There are 423 disfluent sentences in the corpus.

6.1.3 ATIS corpus

The ATIS corpus is a large corpus of human-computer dialog in the air travel planning domain. In the ATIS task subjects were given various scenarios and they had to solve them by interacting with a computer. More details can be found in (Dahl et al., 1994), we discuss here specifically about the disfluent data in the ATIS corpus. The dataset consists of 2586 disfluencies of which 2320 are basic disfluencies while 266 are complex disluencies. There are 1228 disfluent sentences in the corpus.

6.1.4 Disfl-QA dataset

DISFL-QA (Gupta et al., 2021b) extends the current SQUAD-v2 dataset, a question-answering dataset that contains curated Wikipedia paragraphs and accompanying questions. Each query linked with the paragraph was routed to a human annotator and assigned the task of creating a contextual disuency utilising the paragraph as a source of distractions.

7 Data augmentation

Due to the scarcity of gold-standard data for disfluency correction, various data augmentation techniques have been proposed to overcome this challenge. (Yang et al., 2020) propose a *Planner-Generator* based architecture for generating disfluencies from fluent sentences. The Planner determines the optimal location for inserting disfluent segments, and the Generator produces appropriate disfluent segments accordingly.

In (Lee et al., 2020), the authors utilize auxiliary tasks, namely, Named Entity Recognition (NER) and Part-of-speech Tagging (POS), along with disfluency detection to improve the performance of the model. However, since the Switchboard disfluency detection data lacks NER tags, an off-the-shelf model is used to annotate silver-standard NER training data.

To generate large-scale disfluency detection data, (Passali et al., 2022) employ rule-based methods, with a focus on repetition, replacement, and restart disfluencies. Similarly, (Rocholl et al., 2021) fine-tune a BERT model (Devlin et al., 2018) on the Switchboard dataset to predict disfluency labels for the Fisher corpus (Cieri et al., 2004), and then use the automatically labeled silver data as an additional source of training data for further finetuning.

8 Evaluation Metrics

In previous works, disfluency detection models have been evaluated mainly using token-level precision, recall, and f1 score. However, some studies have also explored the use of BLEU score to evaluate the quality of generated fluent sentences by comparing them with reference sentences.

9 Conclusion

In conclusion, the task of distinguishing between reduplication and repetition has been a challenging problem in the field of natural language processing. Researchers have explored different approaches and techniques to identify disfluencies. In addition, various datasets and data augmentation techniques have been explored for disfluency detection, which is an essential task for distinguishing between reduplication and repetition in speech. We also looked at the definition of reduplication and its role in English and Hindi languages.

Despite the progress made in this field, there are still many challenges that need to be addressed, such as developing more accurate and efficient models for disfluency detection, improving the quality and quantity of annotated datasets, and handling various types of disfluencies.

Overall, this survey paper has provided a comprehensive overview of the current literature in disluencies in speech and reduplication in languages, which can be used for the task of distinguishing between reduplication and repetition.

References

Christopher Cieri, David Miller, and Kevin Walker. 2004. The fisher corpus: A resource for the next generations of speech-to-text. In *LREC*, volume 4, pages 69–71.

Deborah A Dahl, Madeleine Bates, Michael K Brown, William M Fisher, Kate Hunicke-Smith, David S Pallett, Christine Pao, Alexander Rudnicky, and Elizabeth Shriberg. 1994. Expanding the scope of the atis task: The atis-3 corpus. In *Human Language Technology: Proceedings of a Workshop held at Plainsboro, New Jersey, March 8-11, 1994*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

John J Godfrey, Edward C Holliman, and Jane McDaniel. 1992. Switchboard: Telephone speech corpus for research and development. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, volume 1, pages 517–520. IEEE Computer Society.

Charles Goodwin. 1986. Between and within: Alternative sequential treatments of continuers and assessments. *Human studies*, 9(2):205–217.

Aditya Gupta, Jiacheng Xu, Shyam Upadhyay, Diyi Yang, and Manaal Faruqui. 2021a. Disfl-qa: A benchmark dataset for understanding disfluencies in question answering. *arXiv preprint arXiv:2106.04016*.

Aditya Gupta, Jiacheng Xu, Shyam Upadhyay, Diyi Yang, and Manaal Faruqui. 2021b. Disfl-QA: A benchmark dataset for understanding disfluencies in question answering. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3309–3319, Online. Association for Computational Linguistics.

Donald Hindle. 1983. Deterministic parsing of syntactic non-fluencies. In *21st Annual Meeting of the Association for Computational Linguistics*, pages 123–128.

Charles F Hockett. 1973. Where the tongue slips, there slip i. *Speech errors as linguistic evidence*, pages 93–119.

Matthias Honal and Tanja Schultz. 2003. Correction of disfluencies in spontaneous speech using a noisy-channel approach. In *Interspeech*.

U.R. Jain. 2007. *Advanced Hindi Grammar*. Center for South Asia Studies, University of California, Berkeley.

Jacqueline C Kowtko and Patti Price. 1989. Data collection and analysis in the air travel planning domain. In *Speech and Natural Language: Proceedings of a Workshop Held at Cape Cod, Massachusetts, October 15-18, 1989*.

Dongyub Lee, Byeongil Ko, Myeong Cheol Shin, Taesun Whang, Daniel Lee, Eun Hwa Kim, EungGyun Kim, and Jaechoon Jo. 2020. Auxiliary sequence labeling tasks for disfluency detection. *arXiv preprint arXiv:2011.04512*.

Willem JM Levelt. 1983. Monitoring and self-repair in speech. *Cognition*, 14(1):41–104.

Annie Montaut. 2009. Reduplication and echo words in hindi/urdu. *Annual review of South Asian languages and linguistics*, pages 21–91.

Tatiana Passali, Thanassis Mavropoulos, Grigorios Tsoumakas, Georgios Meditskos, and Stefanos Vrochidis. 2022. Lard: Large-scale artificial disfluency generation. *arXiv preprint arXiv:2201.05041*.

Pranav Rajpurkar, Robin Jia, and Percy Liang. 2018. Know what you don't know: Unanswerable questions for squad. *arXiv preprint arXiv:1806.03822*.

Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. 2016. Squad: 100,000+ questions for machine comprehension of text. *arXiv preprint arXiv:1606.05250*.

Johann C Rocholl, Vicky Zayats, Daniel D Walker, Noah B Murad, Aaron Schneider, and Daniel J Liebling. 2021. Disfluency detection with unlabeled data and small bert models. *arXiv preprint arXiv:2104.10769*.

Edward Sapir. 2004. *Language: An introduction to the study of speech*. Courier Corporation.

Emanuel A Schegloff. 1987. Recycled turn beginnings: A precise repair mechanism in conversations turn-taking organization. *Talk and social organisation*, 1(1):70–85.

Elizabeth Ellen Shriberg. 1994. *Preliminaries to a theory of speech disfluencies*. Ph.D. thesis, Citeseer.

Jingfeng Yang, Diyi Yang, and Zhaoran Ma. 2020. Planning and generating natural and diverse disfluent texts as augmentation for disfluency detection. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1450–1460.