



HAL
open science

Exploring (dis)fluency patterns of identical repetitions in Romanian spontaneous speech

Oana Niculescu, Maria Candea

► **To cite this version:**

Oana Niculescu, Maria Candea. Exploring (dis)fluency patterns of identical repetitions in Romanian spontaneous speech. Disfluency in Spontaneous Speech (DiSS) Workshop 2023, Aug 2023, Bielefeld, Germany. pp.66-70, 10.21437/diss.2023-14 . hal-04710656

HAL Id: hal-04710656

<https://hal.science/hal-04710656v1>

Submitted on 6 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License



Exploring (dis)fluency patterns of identical repetitions in Romanian spontaneous speech

Oana Niculescu¹, Maria Candea²

¹Romanian Academy Institute of Linguistics "Iorgu Iordan - Al. Rosetti", Romania

²Sorbonne Nouvelle Université, Laboratory Clesthia EA 7345, France

oeniculescu@yahoo.com, maria.candea@sorbonne-nouvelle.fr

Abstract

This paper offers a preliminary analysis of identical repetitions in terms of frequency, distribution and temporal patterns in Romanian spontaneous speech. The study is carried out on a recently developed Romanian speech corpus suited for analyses at the interface between phonetics and laboratory phonology. Six hours of addressed monologues from six adult native speakers (3F/3M) were examined, resulting in 737 repetitions as immediate repeats, with 69% of the data pertaining to male speakers and 31% to female speakers. Our research questions revolve around the complex interplay between identical repetitions (IRs), silent pauses (#) and filler particles (FPs). The results show both cross-linguistic similarities and differences in terms of distribution and duration patterns of IRs, individualizing Romanian in the context of both Romance and Germanic languages. An important contribution of this current study is that it broadens our understanding of (dis)fluencies based on a lesser-documented European language.

Index Terms: disfluent repetitions, duration patterns, corpus linguistics, filler particles, Romanian spontaneous speech

1. Introduction

Analyses of silent pauses in spontaneous speech, as well as different planning, repair and hesitation markers have received a great deal of attention in the last decades, in the frame of very different theoretical or pragmatic approaches like conversational analysis, cognitive science, syntax, automatic language processing. Most of the data were derived especially from well-resourced languages ([1], [2], [3], [4], [5], among other significant studies). However, less is known in this field with respect to an under-resourced Romance language [6] such as Romanian. As a result, our exploratory study aims to showcase Romanian data in order to provide a first reference description of identical repetitions (IRs) in terms of distribution, frequency and temporal patterns.

Due to the fact that representative standard Romanian corpora have been published only in written form [7], [8], making use of annotation systems better suited for pragmatic analyses (ROVA [9], CORV [10], IVLRA [11]), phonetic studies concerning disfluencies and hesitation markers in spontaneous speech, particularly silent pauses and fillers, are rather scarce. Newer corpora such as ROMBAC [12] are also more inclined towards written text data acquisition and processing. The audio files which complement CoRoLa, the largest publicly available corpus of Romanian [13], do not allow users access to the entire recordings, the search option being limited to certain words or lemmas. In brief, to our knowledge, there are no in-depth acoustic studies pertaining to

silent pauses derived from standard Romanian speech corpora. Likewise, phonetic analyses on filler particles derived from Romanian data are still rare. Among the few studies which address this topic ([14], [15], [16]) take on a pragmatic and semantic approach. In this study, we chose to deal with IRs (i.e. immediate and identical self-repeats of spoken material, e.g. [pe pe] ‘on on’) because identifying these phenomena in context allows for an inclusive description of the patterns of silent pauses (#) and filler particles (FPs) often cooccurring in the editing process, thus enabling future inter- and intra- language comparisons. Consequently, the present article would constitute a preliminary analysis of silent pauses and filler particles within IRs.

We grounded this exploratory study on a monological corpus so as to avoid overlap in speech and, at the same time, gather natural addressed speech in a (semi)interactive setting. One of the main objectives of the present analysis is linked to broadening our understanding of formal regularities in immediate self-repetitions based on a lesser-documented European language. Our acoustic, temporal and combinatorial description is placed within a distributional framework, upstream of any syntactic, cognitive or interactional interpretation of these phenomena. Our goal is that these initial results prove useful both to analysts of automatic language processing who consider self-repetitions from the point of view of the disfluency of the predictable chain, and to conversational analysts who interpret them from the point of view of the sequence of interaction.

2. Methodology

2.1. Data

The analysis is carried out on 6 hours of addressed monologues (non-pathological speech) pertaining to 6 adult native speakers, 3 female (166 minutes) and 3 male (178 minutes), ages 33 to 45 ($\mu = 36.8$, $\sigma = 3.84$), sharing the same socio-economic background. All subjects possess higher education. Participants declared themselves as monolingual, representative of the Southern dialect on which standard Romanian is based on.

The monologues share the same three main conversational topics centered around past (memories), present (pursuits), and future activities (forthcoming projects). Participants could move freely from one topic to, always addressing the same female researcher performing the recordings. No participation remuneration was administered. All speakers were highly invested in the task, producing monologues of up to 80 minutes. The average length of a monologue per speaker is 57 minutes (female speakers: $\mu = 55\text{min}$, $\sigma = 13\text{min}$; male speakers: $\mu = 58\text{min}$, $\sigma = 24\text{min}$).

The data used in this current analysis were extracted from Ro-Phon, a larger Romanian speech corpus (12 speakers in total, 6 female and 6 men) recorded, annotated and manually-aligned by the main author as part of her postdoctoral research project (the ROC-lingv project, financed by UEFISCDI (2020 – 2022), [17]). It is important to mention that the corpus was not explicitly designed to elicit repetitions. Instead, the primary reason for developing this linguistic resource was so as to relaunch studies at the interface between phonetics and laboratory phonology within the Romanian national research program. Although the annotation scheme did not closely follow [18], disfluencies such as (non)lexical filler particles, silent pauses, lengthening, identical repetitions, self-corrections and false starts were manually annotated in the transcriptions (through TextGrids in Praat [19]).

2.2. Variables and research questions

In order to explore complex fluency and disfluency [20], [21] patterns of identical repetitions (IRs), we focused our attention on the frequency and nature of the repeated items, but also on silent pauses and filler particles present within the editing phase.

In this study, we use the term ‘silent pauses’ when referring solely to silent intervals of up to 3 seconds (in line with [4]), while the term filler particle is employed for vocalic (nasal) planning and hesitation markers (Rom. *ă(m)* [ə, əm]; *î(m)* [i, im]), without taking into consideration clicks, coughing, laughter or breath pauses.

This analysis is limited to word and phrase repetitions, without including prosodic factors surrounding the context of the repeat. We also do not treat prolonged segments separately, as we view the overall duration of the *reparandum* and the *reparans* in connection with pauses and filler particles.

Due to the lack of phonetic data regarding filler particles in standard Romanian, in this paper we will also offer a brief phonetic account of those fillers present within IRs.

In brief, we address the following research questions in this study:

1. Does Romanian spontaneous speech data follow the same distribution and temporal patterns related to IRs observed in other European languages?
2. What duration patterns are employed by native speakers when they produce immediate repetitions, eventually combined with pauses and fillers, in spontaneous speech production?
3. What type of filler particles, if any, surface in the editing phase of an identical repetition?

A total of 737 repetitions as immediate repeats were manually extracted from the corpus and labeled. The results are discussed in the following section.

3. Results and Discussion

3.1. Distribution and frequency of identical repetitions

Even though participants produced evenly-length monologues, male speakers generated more than twice as many IRs than female speakers. 69% of the 737 occurrences pertain to male and 31% to female speakers.

Similar to previous findings on the topic ([1], [4], [22]), in 98% of the extracted data, the *reparandum* (RM) was repeated only once, while two *repairs* (RR) surfaced only in 2% of the cases. Furthermore, in 80% of the cases, repetitions were single words, while two-word repeats surfaced in 15% of the data, and three-word repeats in 3%. Occurrences leading up to 1% of the data came from four to seven-word repeats.

Taking into account the linguistic format of the repetition, we identified four subtypes:

- (1) *function words* ([ku ku] ‘with with’),
- (2) *lexical words* ([ə: volew # ə: volew] ‘ă volley ă volley’),
- (3) *phrases* ([kred kə kred kə] ‘I think that I think that’), and
- (4) *utterances* ([jo vrɛaw sə zbor # jo vrɛaw sə zbor] ‘I want to fly I want to fly’).

Results show that there is a clear preference for function word IRs (77%, N = 569), followed by lexical words (13%, N = 97), phrases (7%, N = 52), and utterances (3%, N = 19). The results are in line with previous studies on other languages which show that disfluencies manifested by repetitions occur mostly at the beginning of the phrases [1], [4], [22], [23] *interalli*.

In our corpus, 28 different IRs patterns were identified. The five most frequent repeats, making up for 91% of the data, are as follows:

- (1) RM RR e.g., [la: la] ‘at at’ (38%, N = 282),
- (2) RM # RR e.g., [saw # saw] ‘or or’ (36.5% N = 269),
- (3) RM FP # RR e.g., [ji ə # ji] ‘and ă and’ (8%, N = 57),
- (4) RM FP RR e.g., [prin əm prin] ‘through ăm through’ (7%, N = 50),
- (5) RM # FP RR e.g., [sə nu # m: sə nu] ‘not to m not to’ (2.17%, N = 16).

Based on these preliminary results, we observe that identical repetitions where the repair immediately follows the *reparandum* share a similar recurrence with repetitions where the repair is delayed through a silent pause. The following recurrent IRs include a filler particle and/or a silent pause between RM and RR.

3.2. Duration patterns of identical repetitions

Results indicate that IRs range between 213ms (RM RR) to 5741ms (RM FP1 # FP2 # RR), with an average duration of 1052ms ($\sigma = 682$), directly proportional with the number of word-repeats (the average duration of one-word repeat is 941ms, while a seven-word repeat extends to 3812ms). In 90% of the data, the *reparandum* is significantly longer than the repairing segment ($p < 0.001$). On average, RM is 247ms longer than RR.

When the *reparandum* is repeated twice, a t-test showed no significant temporal differences between the first and second repairing segment (μ RR1= 230, $\sigma = 256$ vs μ RR2= 231, $\sigma = 139$; $t = -0.037$, $df = 15$, $p = 0.970$). The scores also revealed no significant differences between the average duration of an identical repetition repaired only once (μ IR repeated once = 1048, $\sigma = 684$) compared to IRs repeated twice (μ IR repeated twice = 1242, $\sigma = 532$; $t = -1.389$, $df = 15$, $p = 0.185$).

Figure 1 illustrates the temporal hierarchy of the first five most frequent IRs from the corpus. Identical repetitions where RR is delayed through a single silent pause show a greater range and outliers compared with other high-recurring immediate repeats.

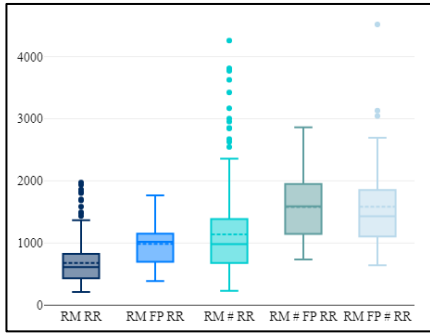


Figure 1: Overall duration (ms) of the five most frequent identical repetitions from the corpus (“RM” = reparandum, “RR” = repair, “FP” = filler particle, “#” = pause).

As summarized in Table 1, duration of IRs was the shortest when the repair immediately follows the *reparandum* and significantly longer when the editing phase contained a filler or a pause. T-test scores revealed there is a statistical difference between the means at the 95% confidence interval in the case of RM RR vs RM FP RR ($p < 0.001$), RM FP RR vs RM # RR ($p = 0.019$), as well as RM RR vs RM # RR ($p < 0.001$). Identical repetitions were the longest when both fillers and silent pauses were present in the editing phase. The scores detected no significant difference in the average duration of IRs depending on the order of elements within the editing phase, i.e. pause – filler vs filler – pause ($t = 0.01$, $df = 28$, $p = 0.986$).

Table 1: Means and Standard Deviations of Raw (ms) and Normalized duration the five most frequent identical repetitions.

	N	Raw Mean	Raw SD	NORM Mean	NORM SD
RM RR	282	679	323	0.26	0.18
RM FP RR	50	985	356	0.43	0.25
RM # RR	269	1138	680	0.22	0.17
RM # FP RR	16	1580	607	0.40	0.28
RM FP # RR	57	1583	710	0.24	0.18

3.3. Duration patterns of silent pauses in identical repetitions

Our results indicate that in 52% of the cases there is at least one silent pause between RM and RR. We noticed a clear preference among speakers to insert only one pause in the editing phase (e.g., [pe # pe] ‘on on’, 383 occurrences), followed by two silent pauses distributed across the IR (e.g., [ku # ə # ku] ‘with ə with’, 19 cases). There was only one identical repetition with three silent pauses in the editing phase present in the corpus (e.g., [de # ə:m # ə # de] ‘of *ăm* ə of’).

In terms of distribution within the five most frequent IRs, when a silent pause occurs, it usually surfaces immediately adjacent to RMs and RRs (269 tokens). Otherwise, silent pauses are adjacent to filler particles (57 cases after, and 16 cases before fillers).

In terms of possible gendered patterns, we observed that male speakers produce more silent pauses compared to female speakers (74%). However, the latter produced longer pauses within the editing phase (male: $\mu = 355$, $\sigma = 305$; female: $\mu = 661$, $\sigma = 576$).

A t-test showed a significant increase in duration ($p < 0.001$) from IRs without a silent pause, having an average duration of 747ms ($\sigma = 367$), to IRs which had a silent pause in the editing phase, averaging at 1334ms ($\sigma = 778$). There is a significant decrease in duration ($t = 2.82$, $df = 25$, $p = 0.009$) of the first silent pause ($\mu = 443$, $\sigma = 422$) compared to the second silent pause ($\mu = 289$, $\sigma = 218$).

Our data revealed different duration patterns of silent pauses depending on the format of the editing phase (Table 2). The overall duration of silent pauses significantly increased in the presence of neighbouring filler particles ($t = -2.304$, $df = 95$, $p = 0.023$). We noticed that post-filler silence duration was slightly longer than pre-filler silence duration. However, this difference is not significant ($t = 0.107$, $df = 50$, $p = 0.915$). This observation is particularly interesting since it places Romanian on an intermediate position between Germanic and Romance languages in terms of the correlation between filler duration and silence duration (for recent cross-linguistic studies on this topic, see [24] for English data, and [25] for German, French and Italian data). These remarks, nonetheless, must be treated with caution, since they involve pauses and filler particles present only in IRs. Future analyses will be carried out by extending to a larger data set with various other contexts.

Table 2: Means and Standard Deviations of Raw (ms) and Normalized duration of silent pauses within the editing phase.

	N	Raw Mean	Raw SD	NORM Mean	NORM SD
RM [#] RR	269	395	384	0.13	0.14
RM FP [#] RR	57	546	565	0.16	0.18
RM [#] FP RR	16	535	285	0.37	0.23

When viewing silent pauses as discrete units of speech, we observed that speakers prefer to simply insert pauses than filler particles in the editing phase (403 cases of pauses compared to 186 instances of filler particles), which suggests that filler particles add an extra pragmatic operation to the simple planning process.

3.4. Duration and acoustic patterns of filler particles in identical repetitions

Even though there are forms of filler particles which are language specific in terms of segmental structure and vocalic timbre ([2], [26], [27], [28], [29], [30], [31]), recent studies have shown that fillers display similar patterns in relation to duration and fundamental frequency [32]. When taking into consideration the possible forms of filler particles in Romanian, it is important to note that the standard language has seven phonemic monophthongs /i, i, u, e, ə, o, a/ and two unary diphthongs /ɛa, oa/ [33, pp. 208]. Among Romance languages, the closed central vowel /i/ is unique [34, pp. 533]. In the UPSID database [35], [36], /i/ appears in 61 out of the 451 languages under survey.

Filler particles in Romanian can take both a vocalic as well as a vocalic-nasal form. Speakers can select from two central vowels /i, ə/ and two nasal sonorants /m, n/. Both the vowel and the nasal can exhibit prolongation.

Contrary to what was observed in a corpus-based approach on German data [37], within IRs derived from Romanian data no glottal filler particle has surfaced. Nonetheless, instances of autonomous glottal filled pauses can be found in the corpus,

suggesting that more investigation is needed on this aspect in future research.

Vocalic fillers are the most frequent form within IRs (129 tokens), followed by nasal fillers (37 tokens – only the bilabial nasal surfaced in the data). Centralized vocalic fillers with a nasal coda have the lowest frequency rate (only 20 tokens).

In terms of potentially gendered patterns, our data point out that male speakers incorporate twice as many filler particles in their IRs compared to female speakers (67%). Moreover, male speakers heavily rely on vocalic fillers (81%), while female speakers use both vocalic (45%) and nasal particles (42%). Considering the low frequency of vocalic-nasal fillers, no statistically significant differences could be derived from the speech corpus.

The data also indicate a clear preference among speakers to insert only one filler particle in the editing phase (168 occurrences), followed by two fillers distributed across the IR (17 cases). There is only one immediate repeat with three filler particles in the editing phase (e.g., [un ə:m # ə un ə] ‘an *ăm* *ă* an *ă*’). A t-test revealed no significant temporal difference between the first and second filler particle from the editing phase ($t = -0.277$, $df = 19$, $p = 0.784$).

With respect to filler particle distribution within IRs, we identified four patterns in association with silent pauses: (1) *concatenated* fillers (i.e. immediately adjacent to RM and/or RR – 74 tokens; e.g., [je: m: je] ‘it is um it is’), (2) *pre-pausal* (72 tokens; e.g., [in i: # in] ‘in *î* in’), (3) *post-pausal* (24 tokens; e.g., [vrəaw sə: # ə:m: vrəaw sə] ‘I want to *ăm* I want to’), and (4) *inter-pausal* filler particles (16 tokens, e.g., [nu: # ə: # nu] ‘no *ă* no’).

With respect to temporal differences among filler particles (Table 3), the results show that vocalic-nasal forms exhibit the highest average duration, being significantly longer than vocalic fillers ($t = 4.859$, $df = 23$, $p < 0.001$), which, in turn, have a longer duration than nasal fillers ($t = 3.238$, $df = 80$, $p = 0.001$).

Our data show that filler particle duration within IRs is consistently shorter than the silent pauses which surround them (observation in line with [38]). In a context such as *RM FP # RR*, the average duration of the filler is 409ms, while the silent pause averages at 546ms. Likewise, in a repetition of the type *RM # FP RR*, the filler particle has a mean duration of 325ms, compared to 535ms average pause duration. Overall, inter- and pre-pausal filler particles register the longest average duration (a t-test revealed no significant temporal differences among the two positions of the fillers, $t = 0.567$, $df = 22$, $p = 0.575$), followed by concatenated and post-pausal fillers (likewise, a t-test showed no significant temporal differences among the two positions of the fillers, $t = 0.417$, $df = 32$, $p = 0.679$). In other words, speakers produce longer filler particles when immediately followed by a silent pause, than when they are immediately followed by speech production (concatenated or post-pausal pre-speech filler particle).

Table 3: Means and Standard Deviations of Raw and Normalized duration of non-lexical filler particles within the editing phase.

	N	Raw Mean	Raw SD	NORM Mean	NORM SD
nasal	37	260	145	0.25	0.22
vocalic	129	357	203	0.28	0.19
vowel + nasal	20	655	262	0.36	0.28

post-pausal	24	301	286	0.2	0.23
concatenated	74	328	211	0.27	0.21
pre-pausal	72	418	208	0.29	0.2
inter-pausal	16	451	215	0.4	0.27

4. Conclusions and outlook

In this study we have examined the complex interplay of identical immediate repeats, silent pauses and filler particles based on Romanian data from the Ro-Phon corpus. Despite the limited size of the corpus and the low number of participants, one important contribution of this pilot study is that it broadens our understanding of fluency and disfluency markers through the observation of duration patterns and distribution of IRs in a lesser-studied Romance language.

In line with the first research question, our data display similarities, to a certain degree, with IRs attested in other languages, namely the high prevalence of one-function word repeat, the longer duration of the *reparandum* in contrast with the repairing segment. The wide range of possible repeats (28 patterns identified in the corpus) sets the data apart from other findings. According to our data, in spontaneous Romanian, identical repetitions where the repair immediately follows the *reparandum* have a similar frequency with repetitions where the repair is delayed through a silent pause.

With respect to the second research question, we observed that IRs duration had the lowest average duration when the repair immediately follows the *reparandum*. In contrast, IRs had the highest average duration when both silent pauses and filler particles surfaced in the editing phase. Silent pauses registered a significant increase in duration in the presence of neighbouring filler particles. However, no significant temporal differences were found when comparing post- and pre-filler silence durations.

In terms of the third research question addressed in this pilot study, we determined that filler particles appear in the editing phase, although not as frequent as silent pauses. Vocalic form particles were the highest frequent fillers present in the data. In terms of segmental structure, vowel + nasal forms had longer average duration than nasal fillers. Overall, autonomous inter-pausal filler particles recorded the highest average duration.

Future work on Romanian identical repetitions can take various directions. Firstly, we intend to examine a larger sampled of recorded data and improve the statistical analyses. Secondly, we aim to further develop the speech corpus by including speakers from other age groups and socio-economic backgrounds in order to allow for an in-depth cross-language comparison of IRs. Future studies could also aim to extend upon the interaction between identical repetitions and other repair types and planning markers.

5. References

- [1] E. Shriberg, *Preliminaries to a theory of speech disfluencies*, Doctoral dissertation, University of California, Berkeley, 1994.
- [2] H. H. Clark and J.E. Fox Tree, “Using Uh and Um in Spontaneous Speaking”, *Cognition*, vol. 84, no. 1, pp.73–111, 2002.
- [3] J. Trouvain, C. Fauth, and B. Möbius, “Breath and Non-breath Pauses in Fluent and Disfluent Phases of German and French L1 and L2 Read Speech”, in *Proceedings of Speech Prosody (SP8)*, Boston, USA, May/June 2016, pp. 31–35.
- [4] M. Candea, *Contribution à l’étude des pauses silencieuses et des phénomènes dits « d’hésitation » en français oral spontané. Étude sur un corpus de récit en classe de français* [Contribution to

- study of silent pauses and “hesitation” phenomena in spontaneous French speech. A corpus study of French oral narratives in class], Unpublished PhD thesis, Université Sorbonne Nouvelle – Paris III, 2000.
- [5] A. Moreno-Sandoval, L. Campillos Llanos, and D. T. Toledano, “A quantitative study of disfluencies in formal, informal and media spontaneous speech in Spanish”, in *Proceedings of IberSPEECH 2012. VII Jornadas en Tecnología del Habla*, Madrid, Spain, Nov. 2012, pp. 164–173.
- [6] D. Trandabăț, E. Irimia, V. B. Mititelu, D. Cristea, and D. Tufiș, *The Romanian language in the digital age*, in *META-NET White Paper Studies*, pp. 1–87, Berlin: Springer, 2012.
- [7] C. Mîrzea Vasile and E. Timotin, “Le roumain contemporain” [Contemporary Romanian], in *Manuel de anthologies, corpus et textes roman*, pp. 612–632, Berlin: De Gruyter Mouton (*Manuals of Romance Linguistics* 7), 2015.
- [8] C. Mîrzea Vasile, “Corpusurile de limba română și importanța lor în realizarea de materiale didactice pentru limba română ca limbă străină” [The Romanian Corpora and their importance in creating teaching materials for Romanian L2], *Romanian Studies Today*, no. 1, pp. 74–95, 2017.
- [9] L. Dascălu-Jinga (coord.), *Româna vorbită actuală (ROVA). Corpus și studii* [Present-day Spoken Romanian. Corpus and studies], Romanian Academy, “Iorgu Iordan – Al. Rosetti” Institute of Linguistics, 2011.
- [10] L. Dascălu-Jinga, *Corpus de română vorbită (CORV). Eșantioane* [Corpus of Spoken Romanian. Samples], București: Oscar Print, 2002.
- [11] L. Ionescu-Ruxândoiu (coord.), *Interacțiunea verbală în limba română actuală. Corpus (selectiv). Schiță de tipologie* [Verbal interaction in present-day Romanian. (Selective) Corpus. Typology scheme], București: Editura Universității din București, 2002.
- [12] R. Ion, E. Irimia, D. Ștefănescu, and D. Tufiș, “ROMBAC: The Romanian Balanced Annotated Corpus”, in *Proceedings of The Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, Istanbul, Turkey, May 2012, pp. 339–344.
- [13] V. Barbu Mititelu, E. Irimia, and D. Tufiș, “CoRoLa – The Reference Corpus of Contemporary Romanian Language”, in *Proceedings of The Ninth International Conference on Language Resources and Evaluation (LREC 2014)*, Reykjavik, Iceland, May 2014, pp. 1235–1239.
- [14] L. Dascălu-Jinga, *Pauzele și întreruperile în conversația românească actuală* [Pauses and Interruptions in Present-day Romanian Conversation], București: Editura Academiei Române, 2006.
- [15] L. Dascălu-Jinga, “Some Romanian non-lexical back channel signals”, *Revue roumaine de linguistique*, vol. 41, no. 3–4, pp. 127–144, 1996.
- [16] A. Avram, “Vocalele neutre ca marcă a ezitării” [Neutral Vowels as Mark of Hesitation], in *Sisteme limbii* [Language Systems], pp. 271–274, București: Editura Academiei Republicii Socialiste Românănia, 1970.
- [17] O. Niculescu, “Developing linguistic resources for Romanian written and spoken language”, in *Proceedings of the 16th International Conference „Linguistic Resources and Tools for Natural Language Processing”*, pp. 21–36, Iași: Editura Universității „Alexandru Ioan Cuza”, 2021.
- [18] J. Hough, L. de Ruiter, S. Betz, and D. Schlangen, “Disfluency and laughter annotation in a light-weight dialogue mark-up protocol”, in *Proceedings of the Seventh International Workshop on Disfluency in Spontaneous Speech (DiSS 2015)*, Edinburgh, UK, Aug. 2015, pp. 49–52.
- [19] P. Boersma, “Praat, a system for doing phonetics by computer”, *Glott International*, vol. 5, no. 9, pp. 341–345, 2001.
- [20] L. Crible, *Discourse markers and (dis)fluency: forms and functions across languages and registers*, vol. 286, John Benjamins, 2018.
- [21] L. Kosmala, *A multimodal contrastive study of (dis)fluency across languages and settings: towards a multidimensional scale of inter-(dis)fluency*, Unpublished PhD thesis, Université Sorbonne Nouvelle, 2021.
- [22] A. Dister, «parler sans accent pour moi c’est sans sans sans bafouiller» Quelles répétitions de formes en français parlé? [Speaking without an accent for me is without without without hesitating. Repetition forms in spoken French], *SHS Web of Conferences*, EDP Sciences, vol. 8, pp. 2017–2031, 2014.
- [23] H. H. Clark and T. Wasow, “Repeating Words in Spontaneous Speech”, *Cognitive Psychology*, vol. 37, no. 3, pp. 201–242, 1998.
- [24] S. Betz and L. Kosmala, “Fill the silence! Basics for modeling hesitation”, in *Proceedings of the Ninth International Workshop on Disfluency in Spontaneous Speech (DiSS 2019)*, Budapest, Hungary, Sept. 2019, pp. 11–14.
- [25] S. Betz, N. Bryhadyr, L. Kosmala, and L. Schettino, “A crosslinguistic study on the interplay of fillers and silences”, in *Proceedings of the Tenth International Workshop on Disfluency in Spontaneous Speech (DiSS 2021)*, Paris 8 University, France, Aug. 2021.
- [26] M. Watanabe, “The constituent complexity and types of fillers in Japanese”, in *Proceedings of The 15th International Congress of Phonetic Sciences*, Barcelona, Spain, Aug. 2003, pp. 2473–2476.
- [27] E. de Leeuw, “Hesitation Markers in English, German, and Dutch”, *Journal of Germanic Linguistics*, vol. 19, no. 2, pp. 85–114, 2007.
- [28] M. Wieling, J. Grieve, G. Bouma, J. Fruehwald, J. Coleman, and M. Liberman, “Variation and change in the use of hesitation markers in Germanic languages”, *Language Dynamics and Change*, vol. 6, no. 2, pp. 199–234, 2016.
- [29] D. Duez, “Acoustico-phonetic Characteristics of Filled Pauses in Spontaneous French Speech: Preliminary Results”, in *Proceedings of the Second International Workshop on Disfluency in Spontaneous Speech (DiSS 2001)*, Edinburgh, Scotland, UK, Aug. 2001, pp. 41–44.
- [30] P. Edmunds, “Buen—buena gente: repair in the Spanish of the South-West”, in *Selected Proceedings of the 8th Hispanic linguistics symposium*, Cascadilla Proceedings Projects. Somerville, MA, pp. 204–213, 2006.
- [31] A. Giannini, “Hesitation phenomena in spontaneous Italian”, in *Proceedings of The 15th International Congress of Phonetic Sciences*, Barcelona, Spain, Aug. 2003, pp. 2653–2656.
- [32] M. Candea, I. Vasilescu, and M. Adda-Decker, “Inter- and intra-language acoustic analysis of autonomous fillers”, in *Proceedings of the Fourth International Workshop on Disfluency in Spontaneous Speech (DiSS05)*, Aix-en-Provence, France, Sept. 2005, pp. 47–51.
- [33] I. Chitoran, *The phonology of Romanian. A constraint-based approach*, Berlin/New York: De Gruyter, 2002.
- [34] M. Renwick, “Romanian”, in T. Meisenburg, C. Gabriel, and R. Gess (eds.) “*Romance phonetics and phonology: an introduction*” *Manual of Romance Phonetics and Phonology* 27. de Gruyter, pp. 531–558, 2021.
- [35] I. Maddieson, *Patterns of Sounds*, Cambridge: Cambridge University Press, 1984.
- [36] I. Maddieson and K. Precoda, “Updating UPSID”, *UCLA Working Papers in Phonetics*, vol. 74, pp. 104–111, 1990.
- [37] M. Belz, “Glottal filled pauses in German”, in *Proceedings of the Eighth International Workshop on Disfluency in Spontaneous Speech (DiSS 2017)*, Stockholm, Sweden, Aug. 2017, pp. 5–8.
- [38] F. Goldman-Eisler, “A comparative study of two hesitation phenomena”, *Language and Speech*, vol. 4, no. 1, pp. 18–26, 1961.