THE INTERACTION OF DISCOURSE MARKERS WITH PROSODIC BOUNDARIES IN ESP

Nicolas Ballier[1], Alice Henderson[2], Adrien Méli[3], Taylor Arnold[4]

[1]Université Paris Cité, CLILLAC-ARP & LLF, F-75013, ,[2]UGA, LIDILEM, [3]Université Paris Cité, CLILLAC-ARP, F-75013, [4]University of Richmond
nicolas.ballier@u-paris.fr, alice.henderson@univ-grenoble-alpes.fr, adrienmeli@gmail.com, taylor.arnold@richmond.edu

ABSTRACT

This paper discusses the acoustic and phonetic realisations of the discourse marker *so* in the EIIDA corpus. We examine how the various functional realisations of *so* are reflected in the F_1/F_2 acoustic space. Using manually annotated paratones and automatic prosodic labels (INTSINT) on the EIIDA corpus, we examine the different correlates for the functional realisations of *so*. We use duration, rhythmic measures, paratone boundaries and INTSINT values as potential cues for the detection of functional realisations.

Keywords: English, discourse markers, paratones, prosodic boundaries, INTSINT.

1. INTRODUCTION

Discourse markers (DMs) are optional, prosodically distinctive "sharing devices and intimacy signals in our everyday talk" [1] and appropriate use demonstrates a speaker's pragmatic and communicative competence (e.g., [2] ; [3]). The functions of DMs have been categorized in different ways, e.g., ideational, textual and/or interpersonal for [4], or for [5] referential, structural, cognitive and/or interpersonal. We are particularly interested in how the interpersonal function surfaces even in well-prepared, non-spontaneous conference talks. Few studies have examined the prosodic environment of DMs in detail [6], and EFL textbooks tend to give only broad guidelines, if any. Corpus studies of authentic use would make it possible to improve guidelines, as some studies have done, e.g., *so* and *well* have been the focus of research. Within the various theoretical and methodological approaches used to study DMs, researchers have examined their prosodic environment in an attempt to correlate prosodic properties with different functions. For example, the prosodic contours of *now* were analysed by [7] to distinguish its sentential and discursive functions and [8] distinguished between semantic and pragmatic connectives, noting that the latter "will often be sentence-initial, followed by a pause and uttered with specific intonation contour".

Our study focused on *so*, for which previous analyses have provided "a fairly scattered picture of [its] functional potential" [4]. Moreover, although all of *so*'s functions occur across diverse text types [2], research on *so* has rarely included prosodic features [6]. As Wennerstrom notes "Because the functions of markers are so broad, any and all analyses of markers [...] can teach us something about their role in discourse"[9].

To that end, we analysed four conference talks by native English speakers for the four aspects studied by Wennerstrom (frequency of use, word position, function, intonational profile), as well as vowel quality. First, we identified the prosodic traits of each token, using manually annotated paratones and automatic INTSINT prosodic labels. Then, we categorized the function of each token as either adverbial or DM by using both lexicogrammatical context (e.g., adverbials retain semantic content and tend not to occur initially), and [10] suggested three conditions for DMs. She posited that they must usually be found in initial position, be sequentially dependent elements which bracket units of talk, and have a range of prosodic contours. Our two research questions are: RQ1: Do the intonational forms or the prosodic profiles of *so* distinguish its adverbial uses from those as a DM? RQ2: Does vowel quality reflect the differences in use? Results are discussed in terms of functional uses of *so*, intonational profile and vowel quality.

The rest of the paper is organised as follows: Section 2 summarises previous research on the analysis of *so* as a discourse marker (DM). Section 3 describes the corpus data and the method of automatic prosodic annotation. Section 4 presents the results and discusses the lexical environment of prosodic boundaries, especially the combinations of discourse markers (lexical bundles). Section 5 discusses vowel quality in relation to speech rate and

coalescence and outlines further research.

2. PREVIOUS RESEARCH

In spite of exploratory analyses of the "organization of conversational interaction" [11], researchers insist on the scarcity of research into the prosody of discourse markers [4]. *So* has been analysed as produced as a single unit, as a stand-alone object in interactional contexts [12]. It has also been analysed in opposition [13] or in combination with other markers. Its interaction with other markers such as *because* [14] has been analysed. [15] and [16] have investigated the combinatorial properties of *so*, demonstrating fewer sequencing constraints than previously claimed for discourse markers. The combining properties of *so* have also been addressed in relation to some constructions. [17] describes the complex emergence of the construction of *so to speak/say* and its interactions with parentheticals. Its hedging function is treated on a par with parentheticals like *as it were* and *if you like*. Different speech situations have been addressed, in dialogical or monological contexts such as seminars [18], in native or non-native speech [3], where it is more frequently used as "I mean" [6].

On the semantics-pragmatics interface, *so* has been analysed as an inferential marker [19] expressing result, inference, connection, topic development or action prompting [20]. In a corpus of video-mediated English as a Lingua Franca, [20] have analysed collocates and frequencies of the uses of *so*, insisting on examples where it is used as 'prompting inference'.

Phonetically, according to the *Longman Pronunciation Dictionary*, "There is an occasional weak form" [21] and the *Cambridge English Pronouncing Dictionary* signals that "the weak form is used only rarely, and only in casual speech before adjectives and adverbs (e.g. *not so bad*)". Our hypothesis is that for discourse marker realisations, the full diphthong /əʊ/ is more likely to be used than in intensive or phraseological uses where schwa /ə/ is more likely. The alternative realisations of *so* when it is not a DM are summed up in [6] as 'an adverb of degree or manner' as 'marking purpose', as 'a part of fixed phrases (e.g., *and so on*)' or 'as a pro-form (e.g., *I think so*)'.

3. MATERIAL AND METHODS

3.1. The EIIDA Corpus

We focus on four lectures (two men, two women) with General American pronunciation models. The speaking style can be defined as corresponding to well-prepared conference talks.

The EIIDA corpus is one of the first multilingual spoken corpora of specialized academic language. It was the main deliverable of a 2012-17 project coordinated by Shirley Carter-Thomas and Jeanne-Marie Debaisieux, from the LATTICE (UMR 8094) research group at the Ecole Normale Supérieure and Université Paris 3 Sorbonne Nouvelle.

EIIDA corpus academic presentations were recorded in different conditions, so all audio files were normalized for sample rate (11,000 Hz), format (wav) and channels (mono) using ffmpeg. Volume normalization across files was carried out following the European Broadcasting Union recommendations, by applying a filter implementing the R128 algorithm.

EIIDA can be used to carry out comparative linguistic analyses on written and spoken academic discourse (research articles vs. conference presentations). It is in three languages - English, French, Spanish - thus facilitating the analysis of the impact of writers' / speakers' linguistic culture on these two modes of communication. Furthermore, as the corpus has written and spoken forms of discourse from geochemistry and linguistics, disciplinary comparisons can be established. The whole spoken corpus corresponds to roughly 20 hours of audio recordings (300,000 tokens). Only recordings by native English speakers were used in the current study. The four manually annotated talks of 22, 34, 28 and 31 minutes are taken from a conference in linguistics (The John Swales Conference). But for one recording, they have a similar speech rate (4.22, 4.24 and 4.26 syllables per second, following [22]).

3.2. Automatic Prosodic Labels

The automatic prosodic labels for pitch were obtained using INTSINT (International Transcription System for Intonation), [23] using SPPAS [24] and a routine to split the talk recordings into smaller files. INTSINT consists in 8 labels representing the annotation: T (Top), M (mid), B (bottom), H (Higher), L (Lower), S (Same), U (Upstepped) and D (Downstepped). To operationalise the detection of rhythmic changes, following De Jong & Wempe's algorithm [22], we extracted syllable peaks during the production of *so* and used their algorithm to count syllable peaks during the production of *so*.

3.3. Prosodic Cues

Following [25], we also took into account paratone boundaries, since paratones (roughly defined in that

study as an increase in pitch range at rhetorical junctures to signal topic shift) correlated significantly with the scores received at an exam for Mandarin speakers of English. For the correlates of paratones in the EIIDA corpus, we followed the results in [26] and in [27], revisiting the correlates postulated for paratones in [28], namely focusing on the acceleration of rhythm at the beginning of paratones. We measured duration, with the hypothesis that uses as DM were more likely to be longer, especially compared to phraseological uses such as *and so on*.

3.4. Hypotheses for Speech Cues of DM uses

For the uses of *so* as DM, we expected occurrences to be longer (duration in ms) and to be closer to diphthong realisations (longer trajectories in plots). Since these occurrences may be in the vicinity of paratone initial boundaries, we also expected them to be closer to paratone boundaries (we computed the mean distance to the preceding paratone boundary) and to verify some of its prosodic correlates (mostly falls in paratone intial contexts). For acceleration, we computed the number of syllable nuclei detected by De Jong and Wempe's algorithm [22] half a second before *so*. For intonation patterns, we expected INTSINT pitch targets to match the frequent Fall pattern identified in [27].

4. RESULTS

4.1. Lexical Bundles

In line with [16], we noticed a strong co-occurrence of *so* with other discourse markers such as *all right* and *OK*.

4.2. Duration

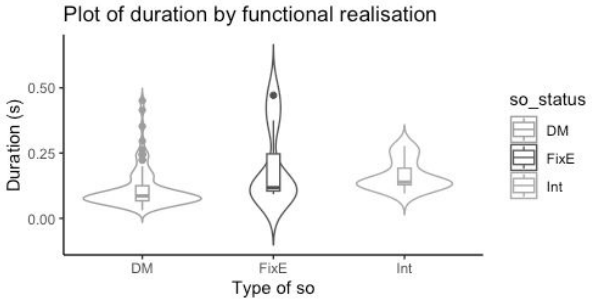

Figure 1: Violin plots of duration for the different realisations of *so* in our EIIDA data

The distribution of duration across types of *so* (see Figure 1) does not correspond to our assumptions for the distribution of duration between DM and Fixed expressions. In other cases, emphatic stress on adverbial *so* lengthened the vowel, so that the duration is not a reliable cue in our data to discriminate DM uses of *so*.

4.3. Acoustic Realisations

Figure 2 plots the diphthong realisations measured at 20% and 80% of the vowel duration of all the occurrences for talk # 2. In this talk, the diphthong is more marked for the intensive uses of *so* (see the black arrow). The other vowel plots similarly show that no complementary distribution between /əʊ/ for DM and /ə/ for adverbial or phraseological intensive uses can be observed. The two European reference

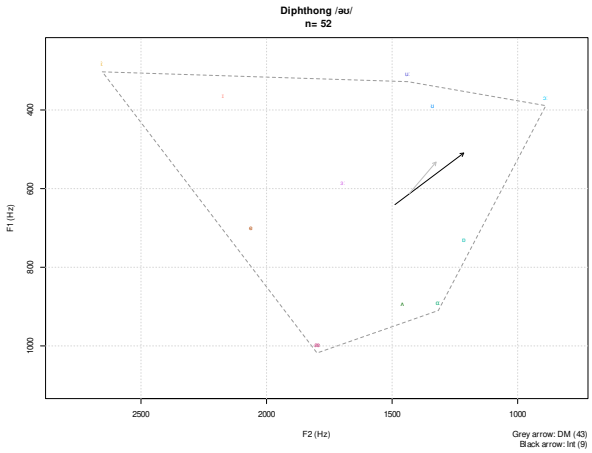

Figure 2: F_1/F_2 values (in Hz) formant plot of *so* in the EIIDA corpus (talk # 02)

pronunciation dictionaries ([21],[29]) do not assign phonetic realisations to categorical realisations for *so*, which they do for other grammatical markers (such as *that*). One way to interpret this result is reference dictionaries are well-advised not to mention categorical distributions for the realisations of *so*.

4.4. Prosodic Cues: INTSINT Labels

With the proviso that our dataset is limited, the INTSINT label associated to the different functional realisations of *so* do seem to indicate a pattern. The low counts and number of INTSINT labels do not lend themselves to parametric tests so we report the raw values for the INTSINT labels in Figure 3.

4.5. Rhythmic Cues: Paratone-like Signals

Taking the INTSINT representation of intonation, DM uses of *so* co-occur with a majority of H and T

	B	D	H	L	M	S	T	U
DM	4	6	9	4	10	3	4	4
FixE	2	1	0	0	0	0	1	3
Int	1	3	1	1	0	0	0	1

Figure 3: Distribution of INTSINT labels for the different realisations of *so* in our EIIDA data

labels, while other uses of *so* mostly co-occur with D , U and L. In other words, DMs tend to be associated with Falls, observed at the beginning of paratones. Not all the files were annotated for paratones, but Figure 4 represents the distance to the preceding initial boundary of the paratone. In spite of outliers, DM uses are earlier in the paratone and, for a fourth of them, used at the beginning of the paratone.

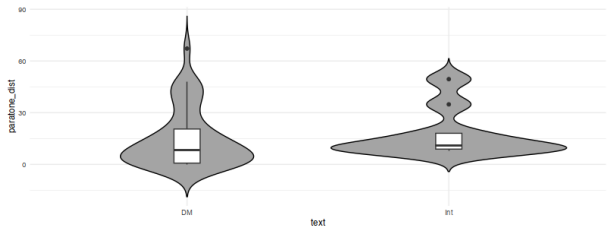

Figure 4: Distance (s) to paratone initial boundaries in talk # 02

5. DISCUSSION

5.1. Connected Speech Processes

While speakers relied sometimes on written notes and speech rate varied from 3.57 syllables per second to 4.26, connected speech processes were observed when investigating the vowel quality of *so*. Several cases of vowel coalescence complexified the analysis for cases like *so at, so we, so actually, so what* and probably account for the duration of some of the phraseological uses (like in *and so on*).

5.2. Prosodic Profiles and Pragmatic Uses

It remains to be seen if more subtle distinctions as listed in [20] are confirmed. They acknowledge no less than 10 pragmatic uses of *so*, some of them are mostly relevant in dialogues such as 'floor-holding device', 'emergence from incipient action', 'Topic developer' and 'Topic sequencer & connection marker Inference'. Some labels could be applied to monological speech such as 'self correction', 'summing up' or 'inferred and explicit result'. Careful manual annotation would be required for further analyses in this direction.

5.3. Further Research

For the detailed semantic interpretations of *so*, [6] notes ambiguous cases in the analysis, which could perhaps be resolved with multimodal cues. Analyzing *so* in multimodal data could be a way to co-cluster criterial features for the detection of DM uses of *so* if specific gestures co-occur with them. A great variability has to be noted as the number of occurrences for *so* ranges from 5 to 53 per lecture in our data, totalling 115 occurrences. These are split into realisations as 95 DM, 10 intensity (adverbial) uses (Int), 7 phraseolocical (fixed) expressions (FixE) and one occurrence as proform. The distribution of the different uses varies across speakers, so that this kind of analysis should be carried out on longer talks (or, for example, on a series of Youtube videos). In any case, the variability of the uses of *so* as a discourse marker is a good candidate to investigate idiosyncrasy. Our prosodic analysis of INTSINT labels could be tested on a larger scale, in a reverse engineering experiment where we would try to predict the status of *so* on the basis of these prosodic properties.

6. CONCLUSION

In this paper, we have analysed the acoustic and prosodic correlates of the uses of *so* in the context of four lectures in linguistics (English for Specific Purposes). For the majority of the occurrences in our data, it seems that prosodic cues are more robust than segmental cues for the automatic detection of the uses of *so*. In his analysis of discourse markers, Zwicky claims that "on the grounds of distribution, prosody, and meaning, discourse markers can be seen to form a class" [30] but does not elaborate on the prosodic properties. It may well be the case that for *so*, prosodic features play a more important role than segmental properties.

7. ACKNOWLEDGEMENTS

We thank the reviewers for comments on an earlier version of this paper. The prosodic annotation of the corpus was funded by a grant to Alice Henderson's lab from CORLI, the French Linguistic Consortium for Corpus, Language and Interaction, via their 2019 call for projects dedicated to corpus delivery (*Appel à Finalisation de corpus*).

8. REFERENCES

[1] Quirk, R., Greenbaum, S., Leech, G., Svartvik, J. 1985. *A comprehensive grammar of the English language*. Longman.

[2] Lam, P. W. 2010. Discourse particles in corpus data and textbooks: The case of well. *Applied Linguistics* 31(2), 260–281.

[3] Müller, S. 2005. *Discourse markers in native and non-native English discourse* volume 138. John Benjamins Publishing.

[4] Buysse, L. 2012. So as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics* 44(13), 1764–1782.

[5] Carter, R., Fung, L. 2007. Discourse markers and spoken english: Native and learner use in pedagogic settings. *Applied Linguistics* 28(3), 410–439.

[6] Buysse, L. 2009. So as a marker of elaboration in native and non-native speech. S. Slembrouck, S., Taverniers, M., Herreweghe, M. V. (eds), *From will to well: Studies in linguistics offered to Anne-Marie Simon-Vandenbergen*. Academia Press, 79–91.

[7] Hirschberg, J., Litman, D. 1993. Empirical studies on the disambiguation of cue phrases. *Computational Linguistics* 19(3), 501–530.

[8] Van Dijk, T. A. 1979. Pragmatic connectives. *Journal of Pragmatics* 3(5), 447–456.

[9] Wennerstrom, A. 2001. *The music of everyday speech: Prosody and discourse analysis*. Oxford University Press.

[10] Schiffrin, D. 1987. *Discourse markers: Studies in interactional sociolinguistics*. Cambridge University Press.

[11] Local, J., Kelly, J. 1986. Projection and 'silences': Notes on phonetic and conversational structure. *Human Studies*, 185–204.

[12] Raymond, G. 2004. Prompting action: The standalone "so" in ordinary conversation. *Research on Language and Social Interaction* 37(2), 185–218.

[13] Lam, P. W. Y. 2010. Toward a functional framework for discourse particles: A comparison of well and so. *Text & Talk* 30(6), 657–677.

[14] Passot, F. 2007. Circularity and discourse progression in conversational English. In: Celle, A., Huart, R. (eds), *Connectives as discourse landmarks* volume 161. John Benjamins Publishing, 117–134.

[15] Koops, C., Lohmann, A. 2013. Discourse marker sequencing and grammaticalization. *Annual Meeting of the Berkeley Linguistics Society* volume 39, 108–122.

[16] Koops, C., Lohmann, A. 2015. A quantitative approach to the grammaticalization of discourse markers: Evidence from their sequencing behavior. *International Journal of Corpus Linguistics* 20(2), 232–259.

[17] Claridge, C. 2013. The evolution of three pragmatic markers: As it were, so to speak/say and if you like. *Journal of Historical Pragmatics* 14(2), 161–184.

[18] Rendle-Short, J. 2003. "So what does this show us?": Analysis of the discourse marker 'so' in seminar talk. *Australian Review of Applied Linguistics* 26(2), 46–62.

[19] Fraser, B. 1999. What are discourse markers? *Journal of Pragmatics* 31(7), 931–952.

[20] Collet, C., Diemer, S., Brunner, M.-L. 2021. So in video-mediated communication in the expanding circle. *World Englishes* 40(4), 594–610.

[21] Wells, J. 2008. *Longman Pronunciation Dictionary*. London: Pearson Longman.

[22] De Jong, N. H., Wempe, T. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41(2), 385–390.

[23] Hirst, D. 2006. Phonetic and phonological annotation of speech prosody. In: Savy, R., Crocco, C. (eds), *(a cura di) Analisi prosodica. teorie, modelli e sistemi di annotazione, Atti del II Convegno Nazionale della Società Italiana di Scienze della Voce (AISV)*. Torriana: EDK, 24-29: Università degli Studi di Salerno 30 novembre-2 dicembre 2005.

[24] Bigi, B. 2015. SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Phonetician* 111-112, 54–69.

[25] Wennerstrom, A. 1998. Intonation as cohesion in academic discourse: A study of chinese speakers of english. *Studies in Second Language Acquisition* 20(1), 1–25.

[26] Méli, A., Ballier, N., Falaise, A., Henderson, A. 2021. An experiment in paratone detection in a prosodically annotated EAP spoken corpus. *Proc. Interspeech 2021*, 2616–2620.

[27] Ballier, N., Méli, A., Arnold, T., Henderson, A. 2022. Revisiting paratone prosodic features with the EIIDA corpus. *Proc. Speech Prosody 2022*, 450–454.

[28] Tench, P. 1996. *The intonation systems of English*. Bloomsbury Publishing.

[29] Jones, D. 2011. *Cambridge English pronouncing dictionary*. Cambridge University Press.

[30] Zwicky, A. M. 1985. Clitics and particles. *Language* 61(2), 283–305.