THE ROLE OF PROSODY IN DISTINGUISHING DIFFERENT TYPES OF CAUSAL RELATIONS IN ENGLISH

Na Hu, Aoju Chen, Hugo Quené, Ted Sanders

Utrecht University, the Netherlands
n.hu@uu.nl, aoju.chen@uu.nl, h.quene@uu.nl, t.j.m.sanders@uu.nl

ABSTRACT

Speakers track how discourse is structured in communication. Lexical cues, e.g. connectives, play an important role in this process as they make discourse relations explicit. However, connectives do not always provide such clear guidance. A case in point is that *because* in English can be used to express both subjective and objective causal relations, leaving the type unspecified. The present study addressed the question as to whether speakers use prosody to encode the differences between causal relations. In a dialogue task, native speakers of American English responded to an interlocutor's questions that elicited either subjective or objective causal relations. Preliminary results show that in comparison with objective causals, subjective causals are produced with higher F_0 maximum, lower F_0 minimum, and also with significantly longer pause between segments. These results suggest that speakers use prosody to distinguish between subjective and objective causal relations in English when lexical cues are absent.

Keywords: causal relations, discourse, subjectivity, prosody, MCMCglmm

1. INTRODUCTION

Causality is a fundamental concept in language and cognition [19, 23]. It can be divided into two types: objective causality and subjective causality, based on the "source[1] of coherence" [17, 26]. Consider for instance the following two sentences (taken from [26]): (1) *Heidi is thrilled because she won the first prize at the art festival;* (2) *Heidi must be talented because she won first prize at the art festival.* In sentence (1), both the consequence and the cause are actual events that happened in the real world, so the causal relation between them is directly observable to others, and thus is objective; while in sentence (2), *Heidi must be talented* is someone's opinion, which is formed by reasoning, based on what has been observed in the reality, so the causal relation between the two segments exists in the mental world, which is not directly observable to others and therefore subjective.

1.1. Lexical cues to subjectivity in causality

These two types of causal relation can be made explicit by lexical cues, such as cue phrases (e.g. *as a result of*) [18], the discourse context [21], implicit causality verbs (e.g. *praise* and *apologize*) [13], syntactic structures [11], and the most commonly used one, connectives. Examples of connectives that point to subjective causality include Dutch *want* [19], French *car* and *puisque* [29], German *denn* [24], and Mandarin *kejian* [14]; and their objective counterparts are Dutch *omdat*, French *parce que*, German *weil*, and Mandarin *yin'er*. Note that English *because* is not on this list, as it can be used to express both subjective and objective causality (see (1) and (2)), leaving their distinctions unspecified on the lexical level [25]. The question that arises at this point is that whether speakers would compensate for this lexical implicitness by means of the non-lexical cue, e.g. prosody, or in other words, whether they would use prosody to distinguish subjective causals and the objective ones in absence of lexical cues. Prosody has a promising role in this respect as it has been shown to be an effective means of disambiguating linguistic ambiguity at various levels, for example, ambiguity in reference [5, 3], in syntactic structure [22], and even in discourse structure [27].

1.2. Prosodic cues to subjectivity in causality

Compared to the amount of knowledge that we have on the lexical cues to subjectivity in causality, much less is known about whether prosody would be used to encode the subjectivity in causality. Researchers working on discourse coherence and connectives have suggested that prosody might play an important role in this respect [20]. Specifically, they have remarked that subjective causal relations consist of two separate propositions, which would formulate two separate intonation units, and therefore require a comma in reading [24]. Objective causals, on the other hand, contain only one proposition, and thus would be pronounced as one integrated intonation unit. However, not working on prosody as their research focus, these researchers did not provide

empirical evidence for their remarks. Being one of the first studies that addressed this issue with real speech data, Couper-Kuhlen [4] has found that subjective causals in English were uttered as two intonation units with the second unit being aligned with a pitch reset at the beginning, whereas objective causals were produced as one intonation units. Also using real speech data, Günthner [7] has made the same observations in German. However, these two studies only based their conclusions on the speculation of pitch contours, but not on the statistic evidences. Related to the current issue, den Ouden et al [15] has explored the correlation between prosody and subjectivity by examining the differences between two groups of discourse relations—subjective (in their terms "pragmatic") relations and objective (in their terms "semantic") relations—in terms of pitch, speech rate, and pause duration. However, the authors did not find any significant differences between these two categories with respect to prosody. One possible explanation to this is that there are too many relations in each category.

1.3. The present study

In the current study, we aim to profile the role of prosody in distinguishing subjective and objective causal relations in English. We collected speech from native speakers of American English through a self-designed question-answer task. This task was carried out in a conversation setting in order to elicit natural utterances. Several acoustic measures, including F_0, speech rate, and pause duration, were extracted to explore whether and how prosody is used to express subjectivity in causality. This paper presents initial findings on the data.

2. METHOD

2.1. Participants

Seventeen native speakers of American English participated in this study. To obtain speech samples that are representative of everyday language use, all speakers were non-actors [12, 10]. Here we report data from nine of the participants (mean age: 25 years, 7 female and 2 male).

2.2. Materials

Experimental items were designed in pairs. There were two items in each pair, one expressing objective causality (3), and the other subjective causality (4). All items were backward causals connected by *because*. For the two items in each pair, only the segments preceding *because* were different (Segment 1, hereafter "S1"), while the segments following *because* were kept identical (Segment 2, hereafter "S2"). In total 15 pairs of items were constructed. In order to prime subjects for different conditions, each item was companied by a short context story (see the italic in (3) and (4)).

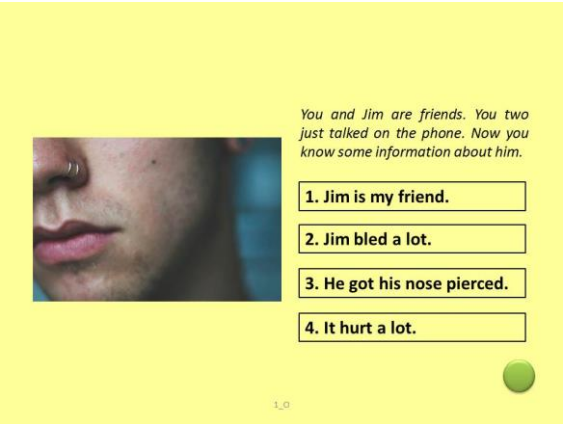
You and Jim are friends. You two just talked on the phone. Now you know some information about him.
(3) [Jim bled a lot]$_{S1}$ [because he got his nose pierced]$_{S2}$.

You and Jim go to the same school, but you don't really know him. You have the following impression about him.
(4) [Jim wants attention]$_{S1}$ [because he got his nose pierced]$_{S2}$.

In order to distract participants from the research purpose, 20 items of concessive relations with *however* were added as fillers. There were also contexts for fillers. Materials were validated by several native speakers of American English.

2.3. The dialogue task

Figure 1: An exemplar PowerPoint slide shown to participants



A dialogue task was designed to elicit naturalistic exemplars of experimental items. During the task, participants first familiarised themselves with information about a character/event from their hypothetical surroundings via PowerPoint slides presented on the screen (one slide per dialogue, see Figure 1 for an example), on which the mini contexts and information fragments were presented in separated "boxes" ordered from 1 to 4. Participants were allowed enough time to read through and digest everything, and then they carried on a conversation with a female experimenter by answering her questions on the character/event. Prior to the task, participants were given instructions on

how they should formulate their answers. They were encouraged to product their answers in a natural way as they would do in daily conversation.

There were always three questions per slide. The target sentences were elicited from the second question, which was *"what do you think of x"* in the subjective condition, or *"what happened to x"* in the objective condition. To answer this question, participants were supposed to combine the segments in box 2 and 3 into one sentence with either *because* or *however*, without changing the order of the "boxes". A first question, being *"who is x"* in most cases, was asked to initiate the conversation. A third question was asked to ensure that target sentences would not be affected by the end-of-conversation prosody. The answers to the first and the third questions were in the first and the third "boxes", respectively.

2.4. Procedure

In order to achieve a relaxed atmosphere during the experiment, the experimenter always conducted casual chats with the participants upon their arrival at the lab. The experiment started with opening remarks by the experimenter, explaining a masked research purpose, in order to give participants a good motivation to participate. After that, participants read instructions on the screen, which were followed by a couple of practice trials to help them get the idea. The experiment started when participants were ready. The experiments were conducted in two blocks, each taking about 20 minutes. Participants could take a short break after the first block. In each session, the participant and the experimenter were seated next to each other by the table, reminiscent of the sitting in a café. The order of the items was randomized in such a way that participants only saw one item in each pair in each block. Moreover, the orders of items were different across participants.

The dialogues were recorded using ZOOM 1 digital recorder (sampling rate of 44.1 kHz, 16 bit, stereo) in a sound-attenuated booth. The recorder was placed 20cm away from the mouth of the participants.

3. ANALYSIS AND RESULTS

In total 270 utterances (30*9 participants) were obtained, two of which were excluded from further analysis due to disfluency. The annotation was conducted in Praat [1] with boundaries of interests being set on different tiers. Pitch contours were manually corrected by deleting outliers.

Acoustic analyses were subsequently conducted using Praat [1]. Three acoustic measures were used to present the static prosodic profile of causal relations that differ in subjectivity, namely, the F_0 maximum (in semitones, relative to 1 Hz), the F_0 minimum (in semitones, relative to 1 Hz), and the speech rate (number of words produced per second). These measures were chosen because they were common measures in discourse prosody studies [15].

Two additional measures were included for analyses, in order to depict the prosodic movement from S1 to S2. These measures were: 1) pause duration between segments; 2) pitch reset, the difference in mean pitch (in semitones, relative to 1 Hz) between S1 offsets and S2 onsets, with mean pitch measured for the first stressed syllable of the first word in S2 and the last syllable of the last word in S1.

Potential effects of subjectivity on prosody were tested in R [16], by fitting a series of MCMCglmm models on the data, using the MCMCglmm package [9]. MCMCglmm has the advantage over other mixed models as it allows more than one dependent variable simultaneously [9], which in our case are the two measures of F_0. Each model included Subjectivity as the fixed effect (two levels: Objective, Subjective), Subject and Item as random effects, and the aforementioned acoustic measures as dependent variables. F_0 maximum and F_0 minimum were modelled together, while speech rate, pause duration, and pitch reset were modelled separately.

3.1. Global features: full-utterance

Subjective causal relations were produced with significantly higher F_0 maximum (post.mean = 0.93 semitone, pMCMC < .001, CI = [0.42, 1.48]) and lower F_0 minimum (post.mean = -1.56 semitone, pMCMC < .001, CI = [-2.51, -0.73]) in comparison with objective causals, suggesting that the overall pitch range was larger in subjective condition than in objective condition. However, the overall speech rate did not differ between Subjectivity conditions (post.mean = -0.06, pMCMC = .728).

3.2. Local features

Table 1: Results for Subjectivity as the predictor of pitch measures (in semitones) for S1 and S2

DepVariable	post.mean	lowerCI	upperCI	pMCMC
S1 F_0Max	0.91	0.31	1.56	0.008**
S1 F_0Min	-1.67	-2.50	-0.69	<0.001***
S2 F_0Max	-0.60	-1.46	0.20	0.168
S2 F_0Min	-1.25	-2.30	-0.03	0.020*

To probe more deeply, we calculated acoustic variations for the two segments separately, trying to locate the effect of subjectivity. Table 1 indicates

that the effects of subjectivity occurred mainly in S1, where subjective causals were produced with higher F_0 Maximum and lower F_0 minimum, in comparison with objective causals. However, neither S1 nor S2 differed significantly in terms of speech rate between conditions.

3.3 Prosodic changes over time

The effect of subjectivity on pause duration was marginally significant: the pause duration preceding *because* was longer in the subjective condition than in the objective condition (post.mean = -0.03 second, pMCMC = 0.054, CI = [-0.002, 0.054]). Durational differences were reported in absolute term because the speech rate did not differ significantly between conditions. However, the pitch difference from the offset of S1 to the onset of S2 did not differ between conditions (post.mean = 0.304 semitone, pMCMC = 0.312, CI = [-0.24, 0.93]).

4. DISCUSSION

The aim of this study was to investigate the role of prosody in distinguishing different types of causal relations. We have found that subjective causals differ significantly from objective causal in terms of F_0 and pause duration: subjective causals were uttered with higher F_0 maximum, lower F_0 minimum, and longer pause duration between segments, in comparison with objective causals.

Relative to objective causals, subjective causals had higher F_0 maximum and lower F_0 minimum, suggesting an expanded pitch range. This is consistent with previous studies on automatic stance (subjectivity) classification: utterances that expressed opinions differed from those expressed facts in terms of pitch [6, 28]. The expanded pitch range in subjective causality suggests that speakers raised the level of effort to engage their interlocutors, in an attempt to get their message across [8]. However, this result should been taken with caution as the differences appear to come primarily from S1 and given the different nature of these clauses—being opinions in the subjective condition and facts in the objective condition—we cannot confidently conclude that this is also true for forward causals, which state facts in the S1 in both subjective and objective causals.

Our finding that subjective causals were generally uttered with prolonged pause between segments confirms the existence of "comma intonation" in subjective causals, as proposed by previous discourse studies [25]. The long pause in subjective causals can be taken as the proof of the effort spent on reasoning in human mind, trying to integrate a claim with its argument. Contrastively, objective causals do not involve such a cognitive process as the relation is already observable in the real world.

However, our data didn't reveal any significant effect of subjectivity on speech rate. This is quite counter-intuitive as we expected that people would speak more slowly in the subjective condition because subjective causals are more complex and therefore require more time to construct than objective causals do, just like in comprehension, where subjective causals require more processing time than objective causals [26, 2]. For the same reason, we expected that people would slow down for the sake of their conversation partners, in order to allow them enough time to process the information. One possible explanation for this discrepancy between the evidence and our expectations is that the speech elicited from the current design was not spontaneous—participants may have already formulated their answers before they actually heard the questions because they could anticipate the upcoming questions after a few trials. So the difficulty in constructing subjective causal relations did not reflect on their speech rate.

5. CONCLUSION

This study aimed to unveil the role of prosody in distinguishing subjective and objective causal relations in English. To achieve this goal, we designed a dialogue task to elicit natural *because*-utterances containing these two types of causal relations. The effects of subjectivity on various acoustic measures were tested by fitting a series of MCMCglmm models. The initial findings include: compared to objective causals, subjective causal relations were uttered with higher F_0 maximum, lower F_0 minimum, and also longer pause preceding *because*.

The results suggest that prosody plays a role in distinguishing different types of causal relations in absence of connective cues. Further work will include other acoustic measures (e.g. final lengthening) into analysis in order to have a more comprehensive understanding of the role of prosody in the production of causal coherence relations.

6. ACKNOWLEDGEMENTS

We would like to thank Jet Hoek for the comments on the dialogue task. We are also grateful to Orlando Jousset and Carla Kist for helping recruit the participants.

7. REFERENCES

[1] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.40, retrieved from http://www.praat.org/.

[2] Canestrelli, A. R., Mak, W. M., Sanders, T. J. M. 2013. Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes* 28, 1394–1413.

[3] Chen, A., Den Os, E., De Ruiter, J. P. 2007. Pitch Accent type matters for online processing of information status: Evidence From natural and synthetic speech. *Linguistic Review* 23, 317–344.

[4] Couper-Kuhlen, E. 1996. Intonation and clause combining in discourse: the case of because. *Pragmatics* 6, 389–426.

[5] Dahan, D., Tanenhaus, M. K., Chambers, C. G. 2002. Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language* 47, 292–314.

[6] Freeman, V., Chan, J., Levow, G. A., Wright, R., Ostendorf, M., Zayats, V. 2014. Manipulating stance and involvement using collaborative tasks: An exploratory comparison. *Proc. 15th INTERSPEECH* Singapore, 2238–2242.

[7] Günthner, S. 1996. From subordination to coordination? Verb-second position in German causal and concessive constructions. *Pragmatics* 6, 323–356.

[8] Gussenhoven, C. 2004. *The phonology of tone and intonation.* Cambridge: Cambridge University Press.

[9] Hadfield, J. D. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software* 33, 21–22.

[10] Hellbernd, N., Sammler, D. 2016. Prosody conveys speaker's intentions: acoustic cues for speech act perception. *Journal of Memory and Language* 88, 70–86.

[11] Hoek, J. 2018. Making sense of discourse: On discourse segmentation and the linguistic marking of coherence relations. Doctoral dissertation, Utrecht University.

[12] Jiang, X., Pell, D. M. 2017. The sound of confidence and doubt. *Speech Communication* 88, 106–126.

[13] Koornneef, A. W., Van Berkum, J. J. A. 2006. On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language* 54, 445–465.

[14] Li, F., Evers-Vermeul, J., Sanders, T. J. M. 2013. Subjectivity and result marking in Mandarin: A corpus-based investigation. *Chinese Language and Discourse* 4, 74–119.

[15] Ouden, H. d., Noordman, L., Terken, J. 2008. Prosodic realizations of global and local structure and rhetorical relations in read aloud news reports. *Speech Communication* 51, 116–129.

[16] R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

[17] Sanders, T. J. M, Spooren, W. P. M. S., Noordman, L. 1992. Toward a taxonomy of coherence relations. *Discourse Processes* 15, 1–35.

[18] Sanders, T. J. M., Sweetser, E. E. 2009. *Causal categories in discourse and cognition.* Berlin: Mouton de Gruyter.

[19] Sanders, T. J. M., Spooren, W. P. M. S. 2015. Causality and subjectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text. *Linguistics* 53, 53–92.

[20] Sanders, T. J. M. & Evers-Vermeul, J. (in press). Subjectivity and causality in discourse and cognition; Evidence from corpus analyses, acquisition and processing. In: Loureda, Ó., Recio Fernández, I.M., Nadal, L., Cruz, A. (eds.), *Methodological Approaches to Discourse Markers.* Amsterdam/Philadelphia: John Benjamins.

[21] Scholman, M. C. J., Demberg, V. 2017. Crowdsourcing discourse interpretations: On the influence of context and the reliability of a connective insertion task. *Proc. 11th Linguistic Annotation Workshop* Valencia, Spain.

[22] Snedeker, J., Trueswell, J. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48, 103–130.

[23] Stukker, N., Sanders, T. J. M., Verhagen, A. 2008. Causality in verbs and in discourse connectives: Converging evidence of cross-level parallels in Dutch linguistic categorization. *Journal of Pragmatics* 40, 1296–1322.

[24] Stukker, N., Sanders, T. J. M. 2012. Subjectivity and prototype structure in causal connectives: A cross-linguistic perspective. *Journal of Pragmatics* 44, 169–190.

[25] Sweetser, E. E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure.* Cambridge: Cambridge University Press.

[26] Traxler, M. J., Bybee, M. D., Pickering, M. J. 1997. Influence of connectives on language comprehension: Eye tracking evidence for incremental interpretation. *The Quarterly Journal of Experimental Psychology Section* 50A, 481–497.

[27] Tyler, J. 2014. Prosody and the Interpretation of Hierarchically Ambiguous Discourse. *Discourse Processes* 51, 656–687.

[28] Ward, N. G., Carlson, J. C., Fuentes, O. 2018. Inferring Stance in News Broadcasts from Prosodic-Feature Configurations. *Computer Speech and Language* 50, 85–104.

[29] Zufferey, S. 2012. 'Car, parce que, puisque' revisited: Three empirical studies on French causal connectives. *Journal of Pragmatics* 44, 138–153.

[1] Various terms were used for this distinction, among them semantic versus pragmatic [17], content versus epistemic [25] and causal versus diagnostic [26].