

Prosodic correlates of subclausal quotation marks

Elke Kasimir

Zentrum für Allgemeine Sprachwissenschaft, Berlin

We present the results of an experimental study which targets prosodic correlates of subclausal quotation marks. We found that written sentences containing passages enclosed by quotation marks were read aloud in a manner that significantly differs in prosody from spoken realizations of corresponding disquoted counterparts. However, we also observed that such prosodic marking of subclausal quotation wasn't strong enough to survive subsequent back-translation into written language: there was no correlation between the presence/absence of quotation marks in the original written examples, and the presence/absence of quotation marks in corresponding back-translations from oral renditions. We investigated three different kinds of uses of quotation marks and found no systematic difference between them with respect to prosodic marking.

1 Introduction

This paper presents an experimental study undertaken at the University of Potsdam in winter 2005/2006 which targets prosodic correlates of quotation marks. Given that written quotation marks in languages like English or German usually have no morphological correlate in spoken language¹, the question has often been raised whether written quotation marks have a counterpart in spoken language in the form of special prosodic marking. The experiments were sought to investigate the extent to which this question can be given a positive answer: to which extent are written quotation marks translated into the spoken idiom by using a recognizably marked prosody, and to which extent does such a marking - if it in fact exists - help to communicate the interpretive differences conveyed by quotation marks? We also asked whether distinctive *uses* of quotation mark behave differently with respect to prosodic marking.

¹ The English form "quote-unquote" and its German counterpart "in Anführungsstrichen" or "in Gänsefüßchen" do not generally replace written quotation marks (see also footnote 7).

The results can be summarized as follows: the presence or absence of subclausal quotation marks in a sentence significantly influences its prosodic realization. However, such prosodic marking fails to be a real spoken substitute for written quotation marks. We conclude this from the fact that quotation marks did not survive translations into the spoken idiom followed by back-translations into the written idiom. Furthermore, the prosodic means were apparently rather diverse and not subject to any specific phonological strategy. Finally, the three different usage patterns of subclausal quotation marks acknowledged in the experimental design (partial reported speech, names or titles, modal uses) were treated alike when being translated into the spoken idiom.

After a review of previous studies on the prosody of quotation and reported speech, the rest of this paper will give an overview of the experimental design, describe the material and methods being used and finally present and discuss the results. A thorough discussion of possible theoretical consequences of the findings reported here will be provided in a separate paper.

2 Previous works on the prosody of quotation

This study is to our knowledge the first investigation ever of the prosody of *subclausal* quotation; it is also the first study which explicitly targets not only reported speech but other uses of quotation marks as well. The prosody of reported speech in spontaneous German and English speech data has been investigated in Br  nner 1991, G  nthner 1997, 1998, Couper-Kuhlen 1998, and Klewitz & Couper-Kuhlen 1999. There is agreement that both direct and indirect reported speech may receive a special prosodic marking. According to Klewitz & Couper-Kuhlen 1999, prosodic means include shifts in global pitch (register), loudness, speech rate or isochronous timing and sometimes also drastic changes in voice quality, for instance whispering. Br  nner 1991 observes that the prosodic marking of reported speech is used to characterize the speaker at hand of stereotypes, often in derogatory or mocking ways (see also Buttny 1997). This suggests that at least certain more articulate forms of prosodic marking are neither a grammatical marker of reported speech, nor, as has often been claimed in the linguistic and philosophical literature, a reproduction of phonological properties of some antecedent utterance, but rather an independent comment on the purported speaker of the utterance. According to Klewitz & Couper-Kuhlen 1999, this kind of mocking or derogatory comment not only occurs with direct reported speech, but also with utterances just describing the quotee, and spoken counterparts of *free indirect style* (*style indirect libre*, *erlebte Rede*) as known from literary language. In the case of direct quotation, prosodic marking doesn't exactly line up with quotation boundaries but is somewhat flexible with regards to onset and offset.

A phonetically oriented study of reported speech based on elicited spontaneous speech data is presented in Bertrand & Espesser 2002. The authors found significant differences in pitch range between direct self-reports on the one hand, direct reports of quotees other than the source speaker on the other hand. They didn't find significant differences between the latter type of direct reported speech and ordinary discourse. Phonetically oriented studies based on corpus data are reported in Grosz & Hirschberg 1992, Hirschberg & Grosz 1992 and Jansen et. al. 2001. Grosz and Hirschberg used spoken material from a professional radio news speaker, together with the original written texts. They investigated how annotators assigned quotational readings when provided either just the sound data (first group), or both sound & text (second group). They found that both groups correlate direct reported speech with greater pitch range and a sentence-final pronounced drop in intensity. Notice that such prosodic marking was produced by a trained speaker and is yet merely a statistical tendency. Jansen et. al. 2001 found that direct reported speech is positively correlated with higher pitch range and the likelihood of the alignment of the report sentence with an intonation phrase boundary.

3 Overview of the experimental procedure

The initial step of our study consisted of a production task where participants translated sentences containing subclausal quotation marks, as well as corresponding sentences lacking the quotation marks, into spoken language. This production task was followed by two classification tasks where participants were asked to match the acoustic tokens from the production task with the corresponding written tokens. In a final reproduction task participants were asked to back-translate items that had been particularly “successful” in the classification tasks back into the written idiom. Figure 1 gives an overview of the overall experimental procedure.

In the classification tasks, things were made as easy as possible for the participants: the latter had just to decide whether a certain passage being embedded in a whole sentence was surrounded by quotation marks, or not. The primary goal of this classification task was to measure the degree to which the presence or absence of quotation marks was encoded in the spoken signal. The participants therefore had a role similar to that of annotators in corpus-based studies. But since this procedure provided us with a measurement of the degree to which particular acoustic tokens were correctly classified, it also helped us to identify particularly "successful" token pairs, that is, tokens pairs where the speakers had encoded the presence or absence of quotation marks in a way that allowed most or all of the listeners to recover this information from the acoustic signal.

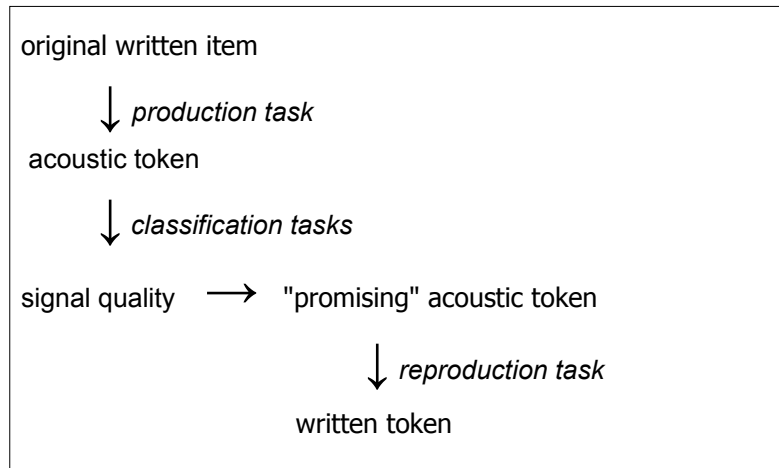


Figure 1: Experimental procedure

Some few of these “successful” items were used for the final reproduction task. This time, participants were asked to listen to whole texts containing either the quote or its disquoted counterpart, and typewrite them into a computer text file. By comparing the results to the original written texts, we could determine to what degree written quotation marks had survived a translation into spoken language and a subsequent back-translation into written language. This in turn seemed to us to be a good indicator of the degree to which the encoding of quotation in spoken language is actually communicatively effective.

4 Material

The material consisted of a collection of subclausal quotation items in Standard German of which the following is an example:

Die Grüne Bundestagsabgeordnete Monika Knoche kritisiert, dass die Verordnung damit einem "faktischen Verbot" von Codein gleichkommt.

The Green deputy Monika Knoche criticises that therewith this bye-law amounts to a "factual verdict" on codeine.

Each of these items were paired with a disquoted variant:

Die Grüne Bundestagsabgeordnete Monika Knoche kritisiert, dass die Verordnung damit einem faktischen Verbot von Codein gleichkommt.

The Green deputy Monika Knoche criticises that therewith this bye-law amounts to a factual verdict on codeine.

For every such sentence pair, a short embedding text was construed such that both variants have a plausible reading in that context.

We used three kinds of subclausal quotation: (I) subclausal direct reported speech (*partial quotes, mixed quotation*), (II) names or titles which were derived from descriptive terms by quotation marks, (III) modal uses of quotation marks, including scare quotes (*modalisierende Verwendungen* in the sense of Klockow 1980):²

The byelaw amounts to a "factual verdict" on codeine.

(= (I) *reported speech*)

Lord Anson's "Journeys round the world" made it into the headlines.³

(= (II) *names or titles*)

The anticlimax was credited to his "political fanaticism".

(= (III) *other / modal uses*)

Six items of each kind were chosen, where each item came in two variants, the first containing, the second lacking the quotation marks in question, yielding a total of 36 texts. All subclausal quotations were basically of the category *noun phrase*. However, in order to facilitate prosodic marking and its subsequent recognition, we let the actual quotation exclude a determiner or adjective whenever this was possible. Furthermore, we took care that the right edge of the quotation didn't coincide with a clause boundary.⁴ All items were somehow derived from naturally occurring examples, mostly newspaper articles and literary texts that were selected from the text corpus database of the Institut für Deutsche Sprache, Mannheim. The surrounding texts were generally taken from the original context but were always simplified both syntactically and with respect to content in order to facilitate loud reading. Some texts had to be invented in order to obtain a context which was compatible with both the quoted and the disquoted version of the expression we were interested in. We took care

² We didn't include metalinguistic quotation (*pure quotation*) as in

(i) *'Boston' is disyllabic.*

because of the obvious difficulties involved in the construction of contexts compatible with both the quoted and the disquoted variant.

³ Notice that in German, names do not receive a special treatment with respect to uppercase-lowercase, so that the German version of the example is disambiguated by quotation marks alone.

⁴ In 14 of 18 items (78%), either a leading determiner or a leading adjective was excluded from the quoted expression. In 17 of 18 items (94%), the right edge of the quoted expression didn't coincide with a phrasal boundary.

that the embedding texts contained further quotation marks in order to conceal the specific differences between the two versions of the texts. The embedding short texts were 65 words on average. The presence/absence of quotation marks generally led to differences in the meaning of the texts, although these differences were in some cases rather subtle.

5 Methods

For the reading task we selected eight speakers of Standard German, four male and four female, who had experience in reading aloud, mostly due to the fact that they regularly read to little children. The participants read all 36 texts first silently and then aloud in four sessions per participant. One session took about 20 minutes. Readings of different versions of the same text were generally at least two weeks apart in order to conceal the difference between the versions. Texts were read again at most one time per session if there were speaking errors in sentences that contained the relevant quotation. Whenever the second reading led to another error, the text was repeated in some subsequent session. The order in which the texts were presented was randomized. The participants didn't become aware of the fact that the experiments had to do with quotation, as was verified by a questionnaire that was presented to the participants at the end of the last sessions.

In the first classification task, 4 participants were asked to rate all 143⁵ acoustic sentence token pairs (a pair consisting of the quoted and the disquoted variant). The audio tracks were presented together with the written counterparts (both in random order) in an Internet browser window, and the participants were asked to determine which track belongs to which sentence version. The participants could listen to the tracks as often as they wanted and had no time limit. Each token pair was presented twice.

In the second classification task, 2 educated participants were asked to rate all 286 acoustic sentence tokens. The audio track was presented with the corresponding written versions (quoted and disquoted) in an Internet browser window and the participants were asked to determine the sentence version to which the audio track belongs. The participants could listen to the track as often as they wanted and had no time limit. Each token was presented once. One listener was an advanced student of phonology, the other one was a researcher in a phonetics scientific project. That way, the increased difficulty of the task, when compared to the first classification, was partially compensated by choosing trained listeners. The goal of the second experiment was to verify that

⁵ One recorded token pair had to be discarded because of a speech error in one of the tokens.

speakers do not just make a relative difference between quoted and disquoted realization, but instead render quotation marks in some absolute, non-relational sense.

As was mentioned earlier, the results of the classification task were also used to select successful items to be used in the subsequent reproduction task. In the above experiment, each token pair can be said to correspond to 12 classifications (two presentations of the pair to 4 participants in the first task, and one classification of one token of the pair to 2 participants in the second task). We randomly selected 6 token pairs which were rated correctly in at least 11 of these 12 classifications for use in the reproduction task.

For the reproduction task, 6 participants were provided with 6 text tokens, each being drawn from a different token pair, in the form of audio files, and were instructed to typewrite the text into a computer file by using the full set of punctuation marks as well as boldface and italics. The listeners were not aware of the goal of the experiment.

6 Results

In both classification tasks, there were more correct than incorrect classifications, indicating that the presence/absence of quotation marks was in fact (prosodically) encoded in the spoken signal. The results are highly significant⁶ (Table 1):

Table 1: Results of the two classification tasks

classifications:	correct	total
1st task (four participants)	763 (67%)	1144
2nd task (two participants)	337 (59%)	572

Furthermore, it turned out that all three kinds of subclausal quotation roughly behaved identical in the classification task (Table 2).

⁶ The post-hoc probability that the high number of correct classifications is accidental is $p < 2 \cdot 10^{-30}$, $p < 8 \cdot 10^{-6}$ respectively, according to the one-sided Sign Test.

Table 2: Kinds of quotation in the 1st classification task

classifications:	correct	total
(I) reported speech	239 (63%)	383
(II) names or titles	267 (70%)	384
(III) other	257 (69%)	377

We finally observed a considerable variability in the performance of "teams" consisting of a speaker and a listener (Figure 2).

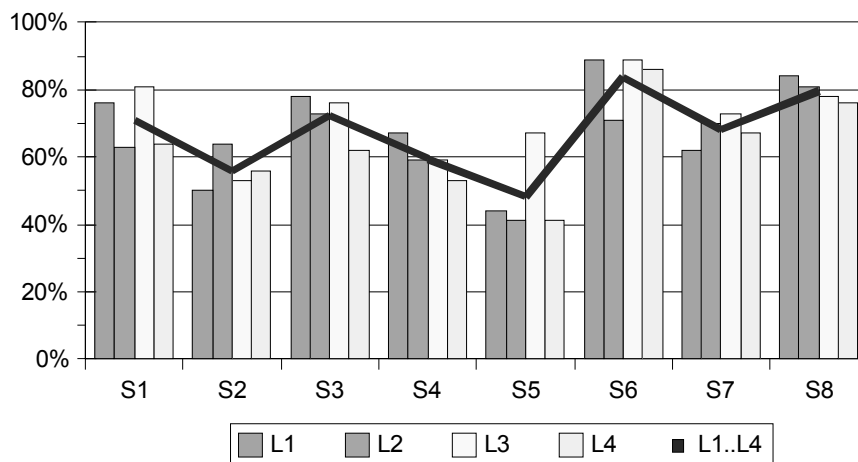


Figure 2: 1st classification task: speakers (S) and listeners (L)

In particular it appears as if the tokens produced by speaker 5 have been merely confusing to listeners 1, 2 and 4, whereas speaker 3 is moderately successful in grasping the quoted-disquoted distinction (see again Figure 2).

In the reproduction experiment we found that even though the tokens being used had been successfully classified in the foregoing classification tasks, not even a tendency towards a reproduction of the quoted-disquoted distinction could be observed (Table 3).

Table 3: Reproduction task

Versions	translations into quoted versions	expected
non-quoted	3	0
quoted	3	18

We were somewhat liberal in what we regarded as a correct reproduction by allowing both quotation marks and italics and by accepting cases where the marking appeared in approximately the same region which was quoted in the original items (see Table 4).

Table 4: List of reproductions that were classified as *quotation* reproductions in the reproduction task. Columns (left to right): the version underlying the acoustic token, the expression which was quoted in the quoted version of the original item, the unedited (!) reproduction as produced by the listener-writer.

Version	Quote	Reproduction
disquoted	Katharina Ehler	An einem Mittwoch verlässt die „Katharina Ela“ <i>Marseille</i> und begibt sich auf die Reise richtung Lateinamerika.
disquoted	Katharina Ehler	An einem Mittwoch verlässt die <i>Katharina Ela</i> Marseille und begibt sich auf die Reise Richtung Lateinamerika.
quoted	Katharina Ehler	An einem Mittwoch verlässt die <i>Katharina-Ela</i> Marseille und begibt sich auf die Reise Richtung Lateinamerika.
quoted	Katharina Ehler	An einem Mittwoch verlässt die „Katharina Ehler“ Marseille und begibt sich auf die Reise Richtung Latein Amerika.
disquoted	Faktischen Verbot	Die Grüne Bundestagsabgeordnete Monika Knochel kritisiert, dass die Verordnung damit „einem praktischen Verbot von Codein gleichkommt“.
quoted	Reise um die Welt	Lord Ansons <i>Reise um die Welt</i> machte im damaligen London Schlagzeilen.

One might finally wonder which particular prosodic means were used in those pairs that were correctly classified in the classification tasks.⁷ Although a phonetic analysis is still missing, our naive auditory impression suggests that at least four different strategies can be distinguished: an audible leading pause, an audible trailing pause, a change in the voice quality (i.e. lengthening of stressed

⁷ None of the speakers used any morphological, lexical or syntactic means as a translation of quotation marks.

vowels) and shifting of the location of pitch accents. The latter strategy can be illustrated at hand of the following item:

Der Ford Scorpio ist ein PKW des "oberen Segments" (obere Mittelklasse oder Oberklasse - je nach Motorisierung und Ausstattung).

The Ford Scorpio is an automobile of the "upper segment" (upper middle class or upper class - depending of motor equipment and configuration).

The main accent of the phrase *Der Ford Scorpio ist ein PKW des oberen Segments* (*The Ford Scorpio is an automobile of the upper segment*) was realized in the quoted version on the adjectival modifier *oberen* (*upper*) by one speaker, whereas the same speaker realized the accent in the default position, i.e. on the modified noun *Segment* (*segment*), in the disquoted version.

7 Discussion

Our results suggest that although speakers prosodically mark renderings of quotation marks, they do not do this in a fashion that effectively supports communication. The speakers also do not seem to rely on some unique strategy for the prosodic marking, and do not differentiate between different usages of quotation marks.

How is this result to be interpreted? One natural explanation is that speakers recognized the need to convey the information contributed by quotation marks when reading aloud the example items. However, spoken language doesn't supplied them with a suitable conventional device to do so; and as a consequence, speakers rely on paralinguistic and non-conventional or at least not fully conventionalized devices. These devices may have been derived from experience with stylized spoken language from mass media, or may have been spontaneously drawn from a set of rather unspecific means to highlight parts of a linguistic utterance or mark boundaries. Alternatively, the failure to communicate subclausal quotation marks could just be a performance error that is induced by the somewhat artificial situation of reading a text aloud instead of producing speech spontaneously. This seems to us however rather unlikely given that the experiment relied on previously selected "successful" encodings and given that other features of the texts were conveyed rather accurately.

The results are consistent with previous results for reported speech as summarized above where it was observed that the quoted-disquoted distinction is marked in a merely statistically significant way by various prosodic means. As expected, our speakers didn't use the more articulate prosodic means which Br  nner (1991) and others observed in spontaneous spoken language (see

section 2 above); if these means are in fact used to characterizing the cited speaker at hand of prototypes, they would have been inappropriate for the newspaper and literary texts that were used in our experiment.

8 References

- Bertrand, R. & R. Espesser (2002). Voice diversity in conversation: a case study. In: *Proceedings of the 1st International Conference on Speech Prosody, Aix-en-Provence, France*, 171-174.
- Brünner, G. (1991). Redewiedergabe in Gesprächen. *Deutsche Sprache* 1, 1-16.
- Buttny, R. (1997). Reported speech in talking race on campus. *Human Communication Research* 23, 477-506.
- Couper-Kuhlen, E. (1998). Coherent voicing. On prosody in conversational reported speech. *Interaction and Linguistic Structures* 1, 1-28.
- Grosz, B. & J. Hirschberg (1992). Some intonational characteristics of discourse structure. In *Proceedings of the International Conference on Spoken Language Processing, Banff*, 429-432.
- Günthner, S. (1999). Polyphony and the "layering of voices" in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech. *Journal of Pragmatics* 31, 685-708.
- Günthner, S. (1997). Direkte und indirekte Rede in Alltagsgesprächen - zur Interaktion von Syntax und Prosodie in der Redewiedergabe. In: P. Schlobinski (ed.). *Zur Syntax des gesprochenen Deutsch*. Westdeutscher Verlag, 227-262.
- Hirschberg, J. & B. Grosz (1992). Intonational features of local and global discourse structure. In: *Speech and Natural Language: Proceedings of a workshop held at Hariman, New York, February 23-26*.
- Jansen, W., J. M. Brenier et. al. (2001). Prosodic correlates of directly reported speech: Evidence from conversational speech. Unpublished paper.
- Klewitz, G. & E. Couper-Kuhlen (1999). 'Quote - Unquote? The role of prosody in the contextualization of reported speech sequences. *Interaction and Linguistic Structure*, 1-34.
- Klockow, R. (1980). *Linguistik der Gänsefüßchen. Untersuchungen zum Gebrauch der Anführungszeichen im gegenwärtigen Deutsch*. Frankfurt a.M.