ANALYSING PROSODY IN SIMULTANEOUS INTERPRETING: DIFFICULTIES AND POSSIBLE SOLUTIONS

Barbara Ahrens
Johannes Gutenberg-University Mainz / FASK Germersheim

1. Introduction

Prosody[1] is an integral part of orally produced texts. Firstly, it is used to structure the acoustic continuum uttered by a speaker and secondly, it is used to give prominence to those parts of the spoken text that the speaker considers to be important. Thus, prosodic elements are essential cues for the listener when processing spoken input (cf. Cutler 1983: 91). Prosody can also be an indicator of the mental-cognitive processes underlying speech production (cf. Goldman-Eisler 1958: 74).

The prosodic feature of intonation – which is defined as pitch movement due to changes of fundamental frequency (F_0) during oral speech production (cf. e.g. Cruttenden 1997[2]: 7, Günther 1999: 62; Schönherr 1997: 12, footnote 4) – has an important role to play in structuring and organizing communicative interaction. Intonation is used to indicate that the speaker will go on speaking or that further elements will follow (cf. Jin 1990: 123ff.; Selting 1995: 50ff.). Intonation has also a social function which depends on the speaker's social status or profession: e.g. priests can easily be distinguished from other professions by the way they speak (cf. Fiukowski and Ptok 1996: 670ff.).

Prosody in bilingual oral communication via an interpreter is as important as in monolingual communicative events. Prosodic elements in the source text (ST) convey meaning that is to be rendered in the target language (cf. Kade 1963: 19), and since the target text (TT) is produced orally, its prosodic features are equally important for the TT's addressee. This aspect was mentioned in early contributions on simultaneous or consecutive interpreting, but was not further developed or considered in interpreting studies for a long time.

2. Prosody in simultaneous interpreting – the state of the art

Interpreters are professional speakers and there is no doubt that their voice and way of speaking are very important (cf. e.g. Alexieva 1990: 5; Cartellieri 1983: 213), especially in the case of simultaneous interpreting (SI) where the

1 Prosody comprises all suprasegmental features that depend on tonal, dynamic and durational parameters. Since these acoustic parameters can be measured objectively, they are important for computer-aided analyses.

interpreter's output is only perceived via headsets. Although most of the few authors who mentioned prosody in SI regard it to be non-marked (cf. Déjean Le Féal 1990: 155; Kirchhoff 1976: 67; Willett 1974: 103), there are also some who put forward a contrary opinion: Barik describes the way simultaneous interpreters speak as "less smooth than 'natural' speech" (Barik 1975: 294) and according to Shlesinger, "the intonational system used in simultaneous interpretation appears to be marked by a set of salient features not found in any other language use." (Shlesinger 1994: 226). For Fiukowski, the unnatural way of speaking in SI is conditioned by the linear ST comprehension and the simultaneous TT planning and production. This means that it cannot be avoided because of the simultaneous interpreting process itself (cf. Fiukowski 1986: 186). Kirchhoff (1976) too acknowledges that difficult speech processing conditions – which are characteristic for SI – can result in a less ideal TT production with hesitations, pauses, etc. (cf. Kirchhoff 1976: 67).

The few studies on prosody in SI conducted so far can be categorized according to the prosodic elements that were examined: pauses, speech rate and segmentation (cf. Alexieva 1988; Barik 1973; Gerver 1969; Goldman-Eisler 1967, 1968; Kreuzpaintner 2001; Lee 1999; Liebig 1994; Shlesinger 1994), accentuation and stress (cf. Pelz 1999; Shlesinger 1994; Williams 1995), as well as intonation and fundamental frequency (cf. Collados Aís 1998; Darò 1990; Shlesinger 1994).

3. Difficulties to be dealt with when analysing prosody in SI

The limited number of studies dedicated to prosody in SI confirms that it has been a neglected field of scientific interest although its importance was acknowledged at the very beginning of interpreting studies (see Section 1). There are several difficulties that have to be dealt with:

- Approaches to analyses and methodology as well as definitions of prosodic phenomena are as diverse as the number of studies. This is not only the case in studies on prosody in SI, but also in studies on prosody in general (for a comprehensive overview with special emphasis to intonation, see Ahrens 2004: 75ff. and 117ff.)
- Purely auditive analyses are subjective, purely automatized speech processing is error-sensitive (cf. Schönherr 1997: 68).
- For a long time, the processing of audio and video data required very powerful computer resources and there was no user-friendly hard- and software available.
- Transcribing and analysing audio and video data is extremely time-consuming, which impairs the processing of large and representative corpora necessary for general conclusions (cf. Gile 1991: 158; Setton 1994: 183).

- Recording professional material in authentic settings is difficult and requires the consent of all parties involved, i.e. interpreters' team, speaker, conference organizer, because of the speaker's and interpreter's copyright for his/her performance (cf. AIIC – General terms of contract; Kalina 1994: 225).
- The scope and objective of the study requires certain quality standards for the recordings, such as sound quality, dual-track recordings etc. (cf. Kalina 1998: 135).
- Any transcription provides a selection of all phenomena comprised in the recordings, i.e. special attention is paid to the elements that are to be analysed (cf. Kalina 1998: 135).
- There are no generally accepted conventions of transcription for prosodic elements.
- Transcribing prosodic phenomena is difficult since they vary a lot. Nevertheless, certain patterns, e.g. falling or rising final pitch movements, can be distinguished and marked in the text (cf. Du Bois *et al*. 1993: 52).

4. Recording the corpus

The corpus used for the study presented here comprised dual-track audio and video recordings of an English ST and three German TTs that had been made in authentic settings. The performance of three parallel booths with two professional interpreters in each of them, all working from English into German, was recorded at the Faculty of Applied Linguistics and Cultural Studies (FASK) of the University of Mainz in Germersheim during a guest lecture on an actual German-English translation job in the field of marketing communication held by a native British English speaker. All the interpreters – four women and two men – were professional interpreters who had been trained either at the FASK, University of Mainz (one woman) or at the University of Heidelberg (the other five) and were working actively and regularly as freelancers on the German private market with an average professional experience of 4.6 years. All six were native speakers of German, two of the women and the two men had English as B language, and the other two women as C language. The four women were working in Booth I and II, the two men in Booth III. In Booth I, there were the two interpreters with English as C language. After the lecture, the six interpreters filled in a questionnaire about their professional background, their preparation, their opinions on the ST and the way the ST speaker had presented it, the problems that had arisen and what they had done to solve them.

The ST speaker, whom the interpreters in the retrospective questionnaires described as "typically British", produced his speech spontaneously, using his manuscript only for short quotations and for planning how to proceed. The ST and the way it was presented can be described as rather informal. For these reasons, the ST showed typical characteristics of spontaneous speech, such as hesitation phenomena, false starts or sudden variations in the speech rate (cf. e.g. Cruttenden 1997[2]: 174; Crystal 1969: 154; Goldman-Eisler 1958: 61).

The audio and video recordings of three parallel booths all interpreting under exactly the same external conditions – same ST, same language pair, same situation, same audience – were made in dual-track quality. For the synchronous video recordings of all three booths and the ST speaker, a fourfold splitscreen was used.

5. Digitizing the data

The ST and the three TTs were digitized by means of *Wavelab 3.0* by Steinberg (see: http://www.steinberg.de/produkte/ps/wavelab/wavelab3/). This software allowed the synchronous digitizing of the two channels of the dual-track recordings. By using a sampling rate of 44.1 KHz (i.e. CD quality) for both channels, it was possible to obtain digitized data with optimum quality. After that, the two channels of each booth were separated and downsampled to 11.025 kHz in order to reduce the quantity of data to an amount that can be handled easily in computer-aided analyses. Due to the synchronous digitizing of the two audio channels of each booth, both channels can be aligned precisely although they were stored in different files (see Section 8).

6. Transcribing the data

The ST and the three TTs were transcribed word by word. Although it is not very reader-friendly to use no interpunctuation, this procedure was chosen in order to avoid a misleading prosodic impression due to commas or full stops. This way of transcribing revealed that prosodic phenomena do not necessarily follow syntactically defined boundaries.

Since the analysis was to focus on pauses, segmentation into intonation units,[2] accentuation patterns and final pitch movements, these phenomena were indicated in the transcriptions. The following conventions of transcription were defined: All texts are written in small letters, syllables in capitals are stressed syllables. 1 line corresponds to 1 intonation unit (IU). "=" indicates lengthening of syllable. "\" means final falling pitch movement, "/" indicates a final rising pitch movement, "—"means final level contour, "¬" indicates a final semi-

2 An intonation unit (IU) is defined as a prosodic unit with a coherent F_0 contour and at least one pitch movement that is perceived as prominent (cf. Ahrens 2004: 111; Huber 1988: 71).

falling pitch movement, "—*" a final rise-level contour. "<0.96>" marks a pause with a duration of 0.96 seconds, "<A>" is a pause due to breathing, "<G>" indicates a pause with noise in the booth, "<M>" means that there is a pause during which the ST speaker is rustling his notes. "{...}" indicates paraverbal comments at the end of the intonation units they refer to, e.g. louder, faster etc. According to these conventions, the beginning of the ST reads as follows:

Example 1 (ST):
THANKS very much/ <1.29>
ehm <1.75A> that's a rather LONG cable here and i have to be CAREful
of— <0.39A>
EHM— <1.69R>
THANK you for the introDUCtion/
THANK you for the invitation as WELL/
i'm very pleased to BE here— <0.61A>
ehm CAN i— <0.13>
before i trip Over the WIRE/ <0.39A>
can i ASK <0.56>
you HOW many of you hea=rd my LECture— <0.39A>
LAST YEAR\
THIS time LAST year— {deeper}
was Anybody here\ <0.41A>
is there Anybody here who was HERE a year ago\ <0.73> {faster}
can you INdicate by raising your HANDS/ <0.69A>
NObody\
THAT'S very GOOD\ <0.18>

7. Calculating speech rate

All perceived pauses were checked and measured using the speech signal of the digitized audio data and included in the transcriptions of all four texts of the corpus. In a second step, all spoken syllables were marked and counted. For each text of the corpus, the total number of syllables was divided by the total length of text (measured in seconds) in order to calculate the average speech rate (measured in syllables per second). The speech rate of the ST corresponds to what is considered a normal speech rate, i.e. an average of 5-8 syllables/second (cf. Goldman-Eisler 1961: 171), although the interpreters said that the ST had been presented very fast.

For calculating the rate of articulation, only the actual speaking time, i.e. total length of text minus total time of pauses, was used. The rate of articulation is in line with Barik's results (1973); he found that the rate of articulation of interpreters is below that of the ST speaker (cf. Barik 1973: 257). Table 1 shows the results of all four texts.

	Total length of text (seconds)	**Number of syllables**	**Speech rate** (syllables/s)	**Total time of pauses** (seconds)	**Rate of articulation** (syllables/s)
ST	4,363.2	16,630	3.81	1,188.0	5.24
TT_I	4,365.6	13,971	3.20	1,347.0	4.63
TT_{II}	4,365.0	14,342	3.28	1,513.8	5.02
TT_{III}	4,363.8	13,856	3.17	1,491.0	4.82

Table 1 Speech rates and rates of articulation of the ST and TT_{I-III}

8. Using PRAAT

For the analysis of the digitized audio data, *PRAAT*, a computer programme especially developed and designed for speech analysis by P. Boersma and D. Weenink at the Phonetic Sciences Department of the University of Amsterdam, was used (for further information, see: http://www.fon.hum.uva.nl/praat/). The dual-track recordings allowed the synchronization of the ST and the respective TT in *PRAAT*.

Figure 1 shows the screen view of a synchronized ST and TT_I paragraph. In each analysis window, the speech signal is reflected in the upper track and the pitch and intensity contour in the lower. One can move forward or backwards in the texts by scrolling to the right or left. It is also possible to view further features in the analysis window, e.g. the spectrogram of both texts which was used for checking word and IU boundaries. Pitch is measured in Hz as indicated on the right-hand side of the analysis windows, intensity in decibels (dB) on the left-hand side. Time (in seconds) is given below the analysis windows. Like in Figure 2 below, further tiers can be defined in order to note down words, syllables, pauses, etc. in time-aligned transcriptions. Any selection of text, sound, speech signal, pitch, etc. can be stored in a time-aligned format in individual files that can be handled more easily for analysis purposes and for generating diagrams.

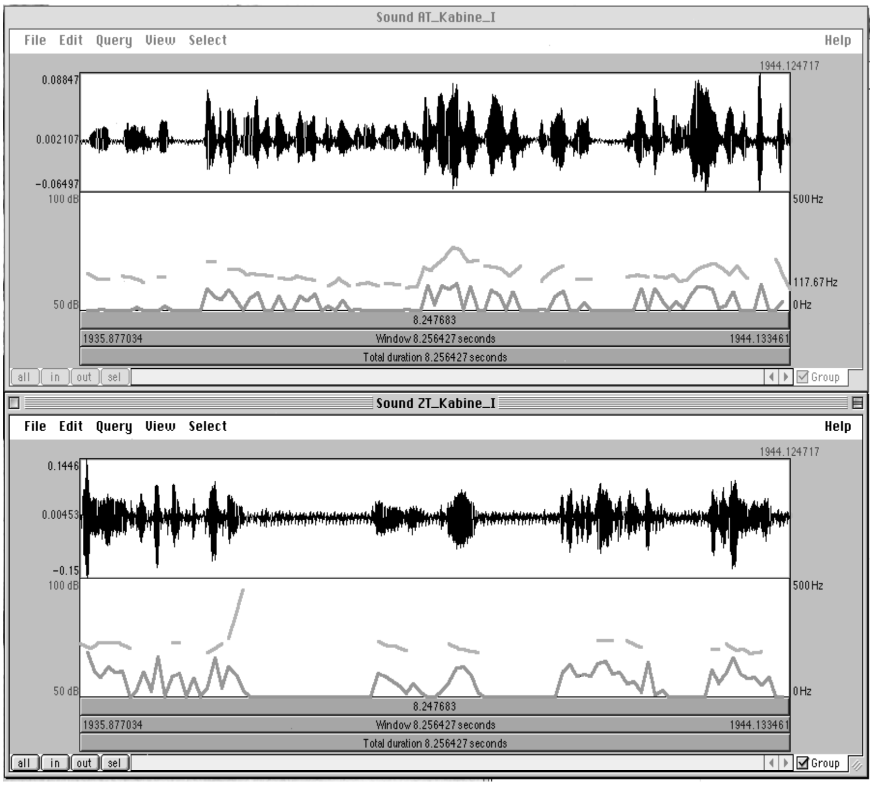

Figure 1 Screen view of a synchronized ST and TT_I paragraph in *PRAAT*

Example 2 (TT_{II}):
ich werde mir auch die proBLEme anschaun—*

The speech signal and the spectrogram were used to define the word boundaries in this IU, the local pitch movement in the word "proBLEme" indicated the accent in it. The characteristic rise-level F_0 contour[3] at the end of this IU can be identified clearly in Figure 2 above and Figure 3 below. Speech signal analysis confirmed these findings (see table 2).

3 A characteristic feature of the TTs was the interpreters' intonational singsong. In this case, the last stressed syllable in an IU showed a rising accent, i.e. the accent was carried out by a salient rising pitch movement. After the stressed syllable, the pitch remained on the frequency that had been reached by the accent until the end of the IU. Since this pattern looks like a combination of rising and level contours, it is called "rise-level" (cf. Ahrens 2004: 209ff.).

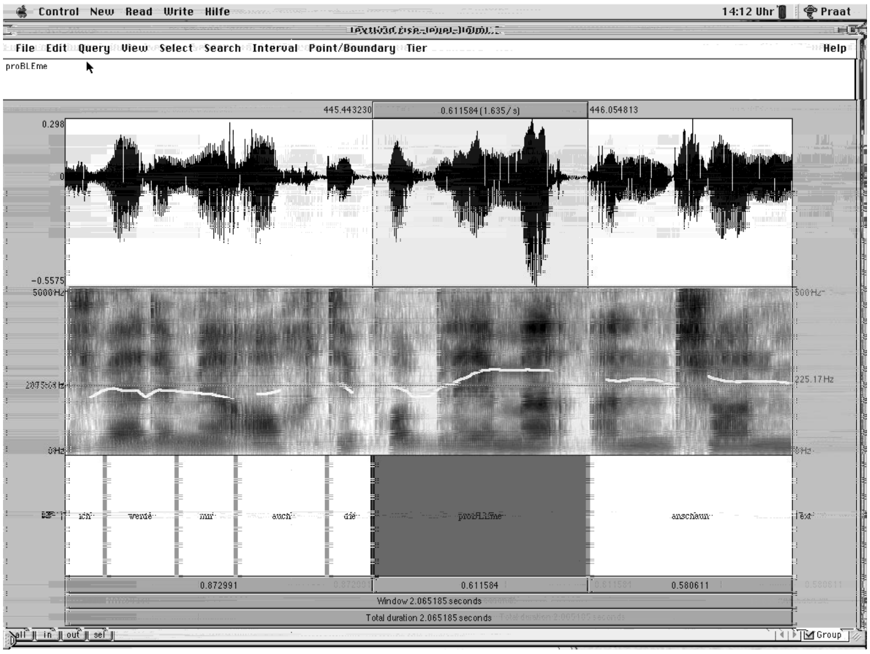

Figure 2 Screen view of example 2 – Speech signal, spectrogram, pitch and text

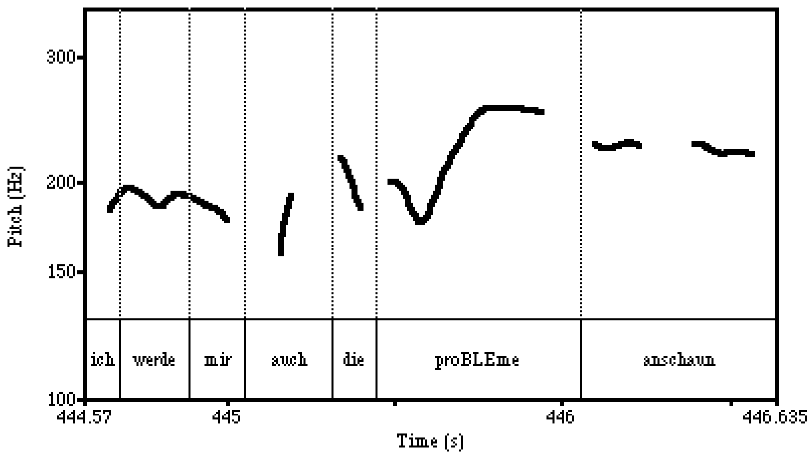

Figure 3 Example 2 - Typical rise-level contour in TT_{II}

	ich werde mir auch die proBLEme anschaun—*
Maximum F_0	253 Hz
Time Maximum F_0	445.79 s
F_0 at the end of IU	219 Hz
Mean F_0 for *level*	233 s

Table 2 Example 2 – Speech signal analysis

9. Analysing segmentation into intonation units

In a first step, the auditively perceived pitch movements were noted down in order to reflect the segmentation into IUs and the accentuation pattern of all four texts. Then, the F_0 contour of all texts was calculated and visualized by means of *PRAAT*. The auditive results of the ST and the TTs were checked against their F_0 contour calculated by the computer. This combined two-step analysis allowed to overcome the shortcomings of purely auditive or purely computerized analyses and helped to obtain a refined picture of the intonational segmentation and the final pitch movement of each IU.

Example 3 (ST):
working with CONnotation and asSOciation\
THIS is the advertisement for ROver\ <0.28>

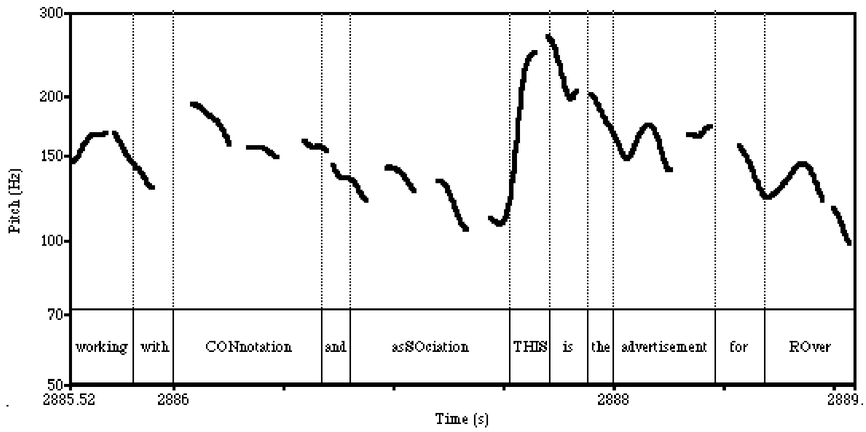

Figure 4 Sequence of two intonation units – Declination and *reset*

In Figure 4, the global falling pitch contour defined as declination (cf. Vaissière 1983: 55 ff.) as well as its *reset* in the second IU are evident. A characteristic of the beginning of a new IU is the *reset* of the intonational contour on the first syllable of the new IU, the so called “onset syllable” (Crystal 1969: 143). By resetting, F_0 returns to the frequency level on which new IUs usually start in an utterance. Also the local pitch movements that are the reason why the words “CONnotation”, “asSOciation”, “THIS” and “ROver” are perceived as being prominent can be seen clearly. These prosodic phenomena were also confirmed by speech signal analysis (see Table 3).

	working with CONnotation and asSOciation	**THIS is the advertisement for ROver\ <0.28>**
Minimum F_0	103 Hz	90 Hz
Time Minimum F_0	2,887.33 s	2,889.07 s
Maximum F_0	187 Hz	279 Hz
Time Maximum F_0	2,886.11 s	2,887.71 s
Mean F_0	146 Hz	162 Hz

Table 3 Example 2 – Speech signal analysis

The continuous acoustic continuum of all three texts was divided into successive IUs. Perceivable as well as measurable boundary markers were: F_0 declination and F_0 reset, characteristic final pitch movements, such as final fall, final rise, final rise-fall, etc. (cf. e.g. Halliday 1966: 117ff., Kohler 1995[2]: 195 ff.), laryngealization, final lengthening (cf. Heuft 1999: 62) and sometimes pauses since they are not a necessary but an additional boundary marker. Very often, the end of an IU is signalized by a combination of several markers, e.g. final lengthening followed by a pause.

10. Conclusion

Fundamental frequency is an objectively measurable parameter for analysis (cf. Gile 2003: 120). For this reason, computer-aided analysis of voice characteristics and prosody helps to gain more insight into the interplay of different prosodic phenomena and its acoustic parameters. Although computer-aided analysis is very helpful, it is always recommendable to cooperate with experts in voice and speaking skills as well as in signal processing. Nevertheless, analysing prosody remains difficult and time-consuming for the researcher. The analysing method applied to the corpus that is presented in this article is a conceptual approach. It parts from the main functions of prosody – structure and prominence – and examines how these manifest themselves in the ST and the TTs (cf. Ahrens 2004: 131ff.). In order to be able to describe prosodic characteristics of simultaneous interpreted texts, these have to be analysed in a first step as if they were monolingual, autonomous texts. In a second step, to be examined is if salient prosodic features relate to the ST or the interpreting process itself (e.g. because of ear-voice-span, hesitations of the ST speaker, waiting for new ST input). An analysis of this kind requires digital or digitized high-quality dual-track recordings.

Hopefully, studies like the one presented here will trigger more research into the most interesting field of prosody in simultaneous interpreting. Further improvements and modifications in the analysing method are necessary and welcome in order to achieve a commonly acknowledged approach to analysing prosody which has not been reached so far.

References

Ahrens B. (2004) *Prosodie beim Simultandolmetschen*, Frankfurt, Peter Lang.

AIIC (s.t.) *Allgemeine Vertragsbedingungen für Konferenzdolmetscher*, Genf, Association Internationale des Interprètes de Conférence.

Alexieva B. (1988) "Analysis of the simultaneous interpreter's output", in P. Nekeman (ed.), *Translation, our future. Proceedings, XIth world congress of FIT*, Maastricht, Euroterm, pp. 484-488.

Alexieva B. (1990) "Creativity in simultaneous interpretation", *Babel* 36/1, pp. 1-6.

Barik H.C. (1973) "Simultaneous interpretation: Temporal and quantitative data", *Language and Speech* 16/3, pp. 237-270.

Barik H.C. (1975) "Simultaneous interpretation: Qualitative and linguistic Data", *Language and Speech* 18/3, pp. 272-297.

Cartellieri C. (1983) "The inescapable dilemma: Quality and/or quantity in interpreting", *Babel* 29/4, pp. 209-213.

Collados Aís Á. (1998) *La evaluación de la calidad en interpretación simultánea: La importancia de la comunicación no verbal*, Granada, Editorial Comares.

Cruttenden A. (1997[2]) *Intonation*, Cambridge, C.U.P., 1st ed. 1986.

Crystal D. (1969) *Prosodic systems and intonation in English*, Cambridge, C.U.P.

Cutler A. (1983) "Speakers' conception of the function of prosody", in A. Cutler and D.R. Ladd (eds.), *Prosody: Models and measurements*, Berlin, Springer, pp. 79-91.

Darò V. (1990) "Voice frequency in languages and simultaneous interpretation", *The Interpreters' Newsletter* 3, pp. 88-92.

Déjean Le Féal K. (1990) "Some thoughts on the evaluation of simultaneous interpretation", in D. Bowen and M. Bowen (eds.), *Interpreting – Yesterday, today and tomorrow*, Binghamton, SUNY, pp. 154-160.

Du Bois J.W., Schuetze-Coburn S., Cumming S. and Paolino D. (1993) "Outline of discourse transcription", in J.A. Edwards and M.D. Lampert (eds.), *Talking data: Transcription and coding in discourse research*, Hillsdale, Lawrence Erlbaum Associates, pp. 45-87.

Fiukowski H. (1986) "Zur Sprecherziehung für Konferenzdolmetscher", *Fremdsprachen* 30/3, pp. 186-191.

Fiukowski H. und Ptok G. (1996) "Sprechstile ausgewählter Berufsgruppen", in H. Kalverkämper und K.-D. Baumann (Hrsg.), *Fachliche Textsorten – Komponenten – Relationen – Strategien*, Tübingen, Gunter Narr, pp. 663-682.

Gerver D. (1969) "The effects of source language presentation rate on the performance of simultaneous conference interpreters", in E. Foulke (ed.), *Proceedings of the Second Louisville Conference on Rate and/or Frequency-Controlled Speech*, Louisville, Center for Rate-Controlled Recordings, University of Louisville, pp. 162-184.

Gile D. (1991) "Methodological aspects of interpretation (and translation) research", *Target* 3/2, pp. 153-174.

Gile D. (2003) "Quality assessment in conference interpreting: Methodological issues", in Á. Collados Aís, M.[a] M. Fernández Sánchez y D. Gile (ed.), *La evaluación de la calidad en interpretación: Investigación. Actas del I congreso internacional sobre evaluación de la calidad en interpretación de conferencias. Almuñécar, 2001*, Granada, Editorial Comares, pp. 109-123.

Goldman-Eisler F. (1958) "Speech analysis and mental processes", *Language and Speech* 1, pp. 59-75.

Goldman-Eisler F. (1961) "The significance of changes in the rate of articulation", *Language and Speech* 4, pp. 171-174.

Goldman-Eisler F. (1967) "Sequential patterns and cognitive processes in speech", *Language and Speech* 10, pp. 122-132.

Goldman-Eisler F. (1968) *Psycholinguistics: Experiments in spontaneous speech*, London and New York, Academic Press.

Günther C. (1999) *Prosodie und Sprachproduktion*, Tübingen, Max Niemeyer.

Halliday M.A.K. (1966) "Intonation systems in English", in A. McIntosh and M.A.K. Halliday, *Patterns of language: Papers in general, descriptive and applied linguistics*, London, Longmans, pp. 111-133.

Heuft B. (1999) "Eine prominenzbasierte Methode zur Prosodieanalyse und -synthese", in W. Hess and W.Lenders (eds.), *Computer Studies in Language and Speech*, Vol. 2, Frankfurt am Main, Peter Lang.

Huber D. (1988) *Aspects of the communicative function of voice in text intonation*, Ph.D. thesis, University of Göteborg.

Jin F. (1990) *Intonation in Gesprächen. Ein Beitrag zur Methode der kontrastiven Intonationsanalyse am Beispiel des Deutschen und Französischen*, Tübingen, Max Niemeyer.

Kade O. (1963) "Der Dolmetschvorgang und die Notation. Bedeutung und Aufgaben der Notiertechnik und des Notiersystems beim konsekutiven Dolmetschen", *Fremdsprachen* 7/1, pp. 12-20.

Kalina S. (1994) "Analyzing interpreters' performance: Methods and problems", in C. Dollerup and A. Lindegaard (eds.), *Teaching translation and interpreting 2. Insights, aims, visions. Selected papers from the second "Language International" Conference, Elsinore, Denmark 4-6 June 1993*, Amsterdam-Philadelphia, John Benjamins, pp. 225-232.

Kalina S. (1998) *Strategische Prozesse beim Dolmetschen: Theoretische Grundlagen, empirische Fallstudien, didaktische Konsequenzen,* Tübingen, Gunter Narr.

Kirchhoff H. (1976) "Das Simultandolmetschen: Interdependenzen der Variablen im Dolmetschprozeß, Dolmetschmodelle und Dolmetschstrategien", in H.W. Drescher und S. Scheffzek (Hrsg.), *Theorie und Praxis des Übersetzens und Dolmetschens*, Frankfurt, Peter Lang, pp. 59-71.

Kohler K.J. (1995[2]) *Einführung in die Phonetik des Deutschen*, Berlin, Erich Schmidt, 1st ed. 1977.

Kreuzpaintner S. (2001) *Der Einsatz von prosodischen Elementen als Strategie beim Simultandolmetschen Eine Fallstudie im Sprachenpaar Spanisch-Deutsch*, Diploma thesis, University of Heidelberg.

Lee T.-H. (1999) "Speech proportion and accuracy in simultaneous interpretation from English into Korean", *Meta* 44/2, pp. 260-267.

Liebig G.-B. (1994) *Der Einsatz von Intonationsstrategien zur Optimierung des Dolmetschergebnisses beim Simultandolmetschen aus dem Englischen ins Deutsche*, Diploma thesis, University of Heidelberg.

Pelz M. (1999) *Betonung in gelesenen Konferenzansprachen und ihre Bedeutung für das Simultandolmetschen: eine akustische Analyse (Englisch-Deutsch),* Diploma thesis, University of Vienna.

Schönherr B. (1997) *Syntax – Prosodie – nonverbale Kommunikation: Empirische Untersuchungen zur Interaktion sprachlicher und parasprachlicher Ausdrucksmittel im Gespräch*, Tübingen, Max Niemeyer.

Selting M. (1995) *Prosodie im Gespräch. Aspekte einer interaktionellen Phonologie der Konversation*, Tübingen, Max Niemeyer.

Setton R. (1994) "Experiments in the application of discourse studies to interpreter training", in C. Dollerup and A. Lindegaard (eds.), *Teaching translation and interpreting 2. Insights, aims, visions. Selected papers from the second "Language International" Conference, Elsinore, Denmark 4-6 June 1993*, Amsterdam-Philadelphia, John Benjamins, pp. 183-198.

Shlesinger M. (1994) "Intonation in the production and perception of simultaneous interpretation", in S. Lambert and B. Moser-Mercer (eds.), *Bridging the gap: Empirical research in simultaneous interpretation*, Amsterdam-Philadelphia, John Benjamins, pp. 225-236.

Vaissière J. (1983) "Language-independent prosodic features", in A. Cutler and. D.R. Ladd (eds.), *Prosody: Models and measurements*, Berlin, Springer, pp. 53-66.

Willett R. (1974) "Die Ausbildung zum Konferenzdolmetscher", in V. Kapp (Hrsg.), *Übersetzer und Dolmetscher: Theoretische Grundlagen, Ausbildung, Berufspraxis*, Heidelberg, Quelle and Meyer, pp. 87-109.

Williams S. (1995) "Observations on anomalous stress in interpreting", *The Translator* 1/1, pp. 47-64.