

The Form and Function of Prosodic Stylization in Spoken Discourse

William Charles Flowe

Universität Konstanz
Philosophische Fakultät
FB Sprachwissenschaft
Januar 2002

in memory
Boyce M. Flowe

Preface

The work on this thesis has been carried out during my employment at the University of Konstanz, Germany and was supported by the DFG in the framework of the Sonderforschungsbereich 511: "Literatur und Anthropologie".

I am very thankful to my advisor, Elizabeth Couper-Kuhlen, for giving me the opportunity to research a subject which I am personally interested in, for her supervision in writing this thesis, and for being part of my thesis examination committee. I am also thankful to Susanne Günthner for her generous support and encouragement and for her agreement to be my second thesis reviewer. Many thanks also go to Hubert Schleichert and Peter Pause for joining my thesis examination committee, and John Local who, on short notice, agreed to be my third thesis reviewer.

For our many informal discussions, colloquia, and data sessions I would like to thank my colleagues Gabriele Klewitz, Dagmar Barth, Christine Gohl, Markus Thumm, Bettina Baron, Susanne Göckel and Martin Hartung. I would also like to thank Wille Schwenkel for his technical assistance. Special thanks go to my closest colleague and good friend Beatrice Szczepek.

I'd also like to thank my family in Virginia and most importantly my wife, Annegret Liebers.

PREFACE	3
DEUTSCHE ZUSAMMENFASSUNG.....	8
ABSTRACT	10
THE FORM AND FUNCTION OF PROSODIC STYLIZATION IN SPOKEN DISCOURSE.....	12
1 INTRODUCTION	12
2 METHODOLOGY AND THEORETICAL BACKGROUND	19
2.1 <i>Interpretive Socio-Linguistics and Contextualization Theory</i>	20
2.2 <i>Conversation Analysis (CA)</i>	21
2.3 <i>Interactional Prosodic Research</i>	22
3 THE DATA	23
CHAPTER 1: A GESTALT APPROACH TO PROSODY – GESTALT PRINCIPLES AND THE GROUPING OF THE SPEECH CONTINUUM.....	26
1 INTRODUCTION	26
1.1 <i>The Intonation Unit</i>	28
1.2 <i>A Gestalt Approach to the Prosodic Grouping of the Speech Continuum</i>	30
2 THE PROXIMITY PRINCIPLE	31
2.1 <i>Proximity Principle in Visual Gestalts</i>	32
2.2 <i>Proximity Principle in Musical Gestalts</i>	32
2.3 <i>Proximity Principle in Prosodic Gestalts</i>	33
3 THE PROXIMITY CORRELATE.....	35
3.1 <i>Proximity Correlate and Visual Gestalts</i>	36
3.2 <i>Proximity Correlate in Musical Gestalts</i>	37
3.3 <i>Proximity Correlate in Prosodic Gestalts</i>	38
4 THE SIMILARITY PRINCIPLE.....	42
4.1 <i>Similarity Principle in Visual Gestalts</i>	42
4.2 <i>Similarity Principle in Musical Gestalts</i>	43
4.3 <i>Similarity Principle in Prosodic Gestalts</i>	43
5 THE DIRECTIONALITY PRINCIPLE	47
5.1 <i>Directionality Principle in Visual Gestalts</i>	47
5.2 <i>Directionality Principle Musical Gestalts</i>	49

5.3 Directionality Principle in Prosodic Gestalts	50
6 COMPOUND GESTALTS	54
6.1 Compound Visual Gestalts	54
6.2 Compound Musical Gestalts.....	55
6.3 Compound Prosodic Gestalts	56
7 NESTED GESTALTS	57
7.1 Visual Nested Gestalt.....	57
7.2 Visual Nested Gestalt.....	58
7.3 Musical Nested Gestalts	59
7.4 Prosodic Nested Gestalts.....	60
8 COMPLEX GESTALTS	65
8.1 Complex Visual Gestalt	66
8.2 Complex Musical Gestalt.....	67
8.3 Complex Prosodic Gestalt	67
9 SUMMARY	70

CHAPTER 2: THE INTERACTIONAL RELEVANCE OF COMPOUND, NESTED AND COMPLEX GESTALTS AT, BELOW, AND ABOVE THE INTONATION UNIT LEVEL.....72

1 INTRODUCTION	72
2 PROSODIC GESTALTS AT THE LEVEL OF THE INTONATION PHRASE.....	74
3 NESTED PROSODIC GESTALTS BELOW THE LEVEL OF THE INTONATION PHRASE	79
4 PROSODIC GESTALTS ABOVE THE LEVEL OF THE INTONATION PHRASE	87
5 SUMMARY	91

CHAPTER 3: A GESTALT APPROACH TO THE HIGHLIGHTING AND STYLIZATION OF PROSODIC STRUCTURES.....92

1 INTRODUCTION	92
1.1 Gestalt Theory and Stylization	93
1.2 Stylization: A Language-in-use Perspective.....	96
2. THE HIGHLIGHTING OF INDIVIDUAL GESTALT PARTS	100
2.1 Highlighting Processes in Visual Gestalts	100
2.2 Highlighting Processes of Reduction and Ornamentation in Musical Gestalts.....	102
2.3 Highlighting Processes in Prosodic Gestalts	105
2.4 Summary	116

3	STYLIZATION IN VISUAL, MUSICAL AND PROSODIC GESTALTS	117
3.1	<i>Adjustment, Reduction, and Ornamentation in the Stylization of Visual Gestalts ...</i>	118
3.2	<i>Gestalt Reduction in Music.....</i>	122
3.3	<i>Gestalt Reduction in Prosodic Gestalts.....</i>	125
4.	STYLIZATION IN VISUAL, MUSICAL, PROSODIC GESTALTS: GESTALT ORNAMENTATION	134
4.1	<i>Global Gestalt Ornamentation in Visual Gestalts.....</i>	135
4.2	<i>Global Gestalt Ornamentation in Musical Gestalts.....</i>	136
4.3	<i>Global Gestalt Ornamentation in Prosodic Gestalts</i>	138
5	SUMMARY	140
CHAPTER 4: THE PLACEMENT AND STYLIZATION OF OH IN INFORMING SEQUENCES.....		142
1	INTRODUCTION	142
2	NON-OVERLAPPED OH	146
2.1	<i>Freestanding oh.....</i>	147
2.2	<i>Oh followed by additional turn components.....</i>	153
2.3	<i>Deviant Cases</i>	156
3	NESTED OH GESTALTS.....	159
4	COMPLEX OH GESTALTS.....	164
5	CONCLUSION	166
CHAPTER 5: PROSODIC STYLIZATION IN ASSESSMENT SEQUENCES		168
1	INTRODUCTION	168
2	PREVIOUS WORK ON ASSESSMENTS	170
2.1	<i>Pomerantz (1975, 1984).....</i>	171
2.2	<i>Auer/Uhmann (1982).....</i>	174
2.3	<i>Goodwin/Goodwin (1987, 1992).....</i>	175
2.4	<i>Uhmann (1996).....</i>	176
2.5	<i>Auer/Couper-Kuhlen/Müller (1999).....</i>	177
3	UPGRADES/DOWNGRADES THROUGH PROSODIC STYLIZATION	179
4	PROSODIC STYLIZATION IN ASSESSMENT ACTIVITIES.....	185
4.1	<i>Collaborative Prosodic Stylization in Assessments.....</i>	185
4.2	<i>Prosodic Stylization in Assessments in a Playful Modality: an Independent Gloss of Preference.....</i>	193
5	SUMMARY	200

CONCLUSION	201
APPENDIX: BASIC TRANSCRIPTION CONVENTIONS.....	203
BIBLIOGRAPHY	205

Deutsche Zusammenfassung

Die dieser Arbeit zugrundeliegende Aufgabenstellung entstammt dem Forschungsvorhaben des Teilprojekts 8 „Form und Funktion prosodischer Stilisierungen in kommunikativen Gattungen“ innerhalb der 1999 begonnenen Antragsperiode des Sonderforschungsbereichs 511 „Literatur und Anthropologie“. Die Ziele dieses Teilprojekts, die in der vorliegenden Dissertation behandelt werden, teilen sich in zwei Arbeitsbereiche: 1) formal beschreibende Fragestellungen, 2) funktionale Fragestellungen.

Im ersten Teil der Arbeit werden die formal beschreibenden Fragestellungen behandelt. Ausgehend von der Gestalttheorie, wie sie von Wissenschaftlern wie Christian von Ehrenfels, Kurt Koffka, Wolfgang Köhler, Max Wertheimer, u.a. entwickelt wurde, werden „stilisierte“ prosodische Gestalten untersucht. Es wird behauptet, dass prosodische Gestalten den gleichen Wahrnehmungsprinzipien unterliegen wie visuelle und musikalische Gestalten. Die von Max Wertheimer beschriebenen Hauptprinzipien sind folgende: 1) Prinzip der Ähnlichkeit bzw. der Gleichheit, 2) Prinzip der Nähe, 3) Prinzip der guten Kurve bzw. des gemeinsamen Schicksals.

Diese Prinzipien leiten die Wahrnehmung komplexer prosodischer Gestaltmuster beim Zuhörer. Solche Gestaltmuster werden in der vorliegenden Arbeit „*compound*“, „*nested*“, und „*complex*“ genannt und können größer, kleiner oder genauso groß wie die Intonationseinheit der traditionellen prosodischen Forschung sein.

„*Compound*“, „*nested*“ und „*complex*“ Gestaltmuster sind nicht nur Analytikerkonstrukte, sondern haben für die Gesprächsteilnehmer selbst eine interaktionale Relevanz, die durch eine Orientierung der Interaktanten an den Phänomenen innerhalb der Interaktion gezeigt wird.

Um eine erhöhte Formgebung zu beschreiben, schlägt die vorliegende Arbeit zwei Arten von Gestaltmustern vor: 1) „*Highlighted*“ Gestaltmuster und 2) „*Stilisierte*“ Gestaltmuster. Ein „*highlighted*“ Gestaltmuster beschreibt einen kleinen Teil einer

größeren Gestalt, der erhöht wird. Die prosodische Erhöhung eines längeren Sprachsegmentes wird „stilisiert“ genannt. „*Highlighted*“ und „stilisierte“ Gestaltmuster unterliegen drei Prozessen, die diese Dissertation 1) „*Reduction*“ 2) „*Adjustment*“ und 3) „*Ornamentation*“ nennt. Wie bei den „*compound*“, „*nested*“, und „*complex*“ Gestaltmuster sind auch „*highlighted*“ und „stilisierte“ Gestaltmuster relevant für die Interaktion. Diese Relevanz wird auch durch die Orientierung der Interaktanten an den „*highlighted*“ und „stilisierten“ Gestaltmustern gezeigt.

Im zweiten Teil der Arbeit werden die behaupteten Phänomene funktional beschrieben. Prosodische „Stilisierung“ wird als konversationelle *Practice* betrachtet, die eine wesentliche Rolle in bestimmten Diskursaktivitäten spielt. Die Arbeit untersucht zwei Aktivitäten: 1) *Informing*-Sequenzen, 2) *Assessment*-Sequenzen.

Innerhalb solcher *Informing*-Sequenzen spielen prosodische „Stilisierung“ und Gestaltmuster eine wesentliche Rolle in der Aushandlung des Rederechts. In *Assessment*-Sequenzen können stilisierte prosodische Gestaltmuster benutzt werden, um auf *Upgrades* und *Downgrades* hinzudeuten. Dies geschieht auch dann, wenn die lexikalische Wortwahl der Bewertung dieselbe ist.

Der zweite Teil der Arbeit zeigt somit, dass die im ersten Teil beschriebenen Gestaltphänomene eine wesentliche Rolle in alltäglicher Interaktion spielen.

Abstract

The research for this dissertation was carried out as part of the project „Form and Function of Prosodic Stylization in Communicative Genres“. This project is situated within the interdisciplinary SFB 511 “Literatur und Anthropologie“. The aims of the project dealt with in the dissertation may be divided into two areas: the form of “stylization”, and the function of such “stylization” in interaction.

The first part of the dissertation aims at describing the form of prosodic stylization. Based on gestalt theory as proposed by scientists such as Christian von Ehrenfels, Kurt Koffka, Wolfgang Köhler, and Max Wertheimer, the dissertation considers “stylized” prosodic gestalts. It asserts that prosodic gestalts obey the same principles of perception as do visual and musical gestalts. The main principles of gestalt perception as described by Max Wertheimer are 1) the principle of similarity, 2) the principle of proximity, and 3) the principle of directionality.

These principles guide the listener’s perception of intricate prosodic gestalt patterns. The dissertation terms such gestalt patterns “compound”, “nested”, and “complex”. They can be larger than, smaller than, or the same size as the intonation unit of traditional prosodic research.

“Compound”, “nested” and “complex” gestalt patterns are not mere analytical constructs, but have interactional relevance for the interactants. The interactional relevance is shown through an orientation to the phenomena by the conversation participants within the interaction.

In order to describe such heightened forms this dissertation proposes two kinds of gestalt patterns: “highlighted” gestalt patterns, and “stylized” gestalt patterns. A “highlighted” gestalt pattern describes a small part of a larger gestalt, where this small part displays a heightened form. “Stylization” describes a larger gestalt segment which displays prosodic heightening. “Highlighted” and “stylized” gestalt patterns are influenced by three gestalt processes which this dissertation terms 1) “reduction”, 2) “adjustment” and 3) “ornamentation”. As with “compound”, “nested”

and “complex” gestalt patterns, “highlighted” and “stylized” gestalt patterns are also relevant for the interactants. This relevance is similarly shown through the orientation of the interactants towards the proposed phenomena.

In the second part of the dissertation the function of the proposed phenomena will be described. Prosodic “stylization” is considered a conversational practice which plays an essential role in certain discourse activities. Two kinds of discourse activities are considered here: 1) informings and 2) assessments.

Within informing sequences, prosodic “stylization” and gestalt patterns play an essential role in negotiating turn taking and floor holding rights. In assessment sequences, “highlighted” and “stylized” prosodic gestalt patterns may be used to indicate upgrades and downgrades. This happens even when the lexical choice of the second assessment is the same as the first.

The second part of the dissertation thus shows that the gestalt phenomena described in the first part play an essential role in everyday interaction.

The Form and Function of Prosodic Stylization in Spoken Discourse¹

1 Introduction

All passionate language does of itself become musical – with a finer music than the mere accent; the speech of a man even in zealous anger becomes a chant, a song.

Thomas Carlyle

As an animal species, the human being is a singing creature, but he combines ideas with the musical sounds involved.

Wilhelm von Humboldt

Poetry proper is never merely a higher mode (*melos*) of everyday language. It is rather the reverse (...).

Martin Heidegger

(In Chatwin, 1980)

Often in spoken discourse speakers' utterances become highly musical, a phenomenon observed not only by the Scottish essayist Thomas Carlyle, the

¹ The following work was financed through the *Deutsche Forschungsgemeinschaft*, a central public funding organization for academic research. My topic, "*Form und Funktion prosodischer Stilisierungen in Kommunikativen Gattungen*" ("Form and Function of Prosodic Stylization in Communicative Genres"), is situated within the project, "*Ästhetische Phänomene in mündlichen, kommunikativen Formen und Gattungen: Von Rahmung zu Performance*" ("Aesthetic Phenomena in Spoken Communicative Forms and Genres: from Framing to Performance"), which is again situated within the larger research project: *Sonderforschungsbereich 511 "Literatur und Anthropologie"* ("Literature and Anthropology").

German philologist Wilhelm von Humboldt and the German philosopher Martin Heidegger above, but also intuitively recognized and exploited by interactants themselves in everyday spoken discourse.

For a concrete example, consider the following excerpt from a late night German television talk show. The speaker, Harald Schmidt, is in the middle of a comic monolog about Red Cross ambulance drivers who take advantage of their position in order to run stoplights:

“Schön gesagt”

1 H: einzelne
2 einzelne
3 schwarze
4 schafe im an-
5 sonst
6 sensationellen
7 fantastischen
8 roten
9 .hhh
10 kreuz
11 fhuh das habe ich schön gesagt

1 H: *single*
2 *single*
3 *black*
4 *sheep in the*
5 *otherwise*
6 *sensational*
7 *fantastic*
8 *red*
9 *.hhh*
10 *cross*
11 *fhuh I said that beautifully*

The humorist Schmidt is undoubtedly exploiting the ability of language to become musical in this excerpt, and indeed, the language of music readily adapts itself to describe the excerpt: it is rhythmic, melodic, assonant, alliterative, and as Schmidt himself asserts, “schön” (line 11).

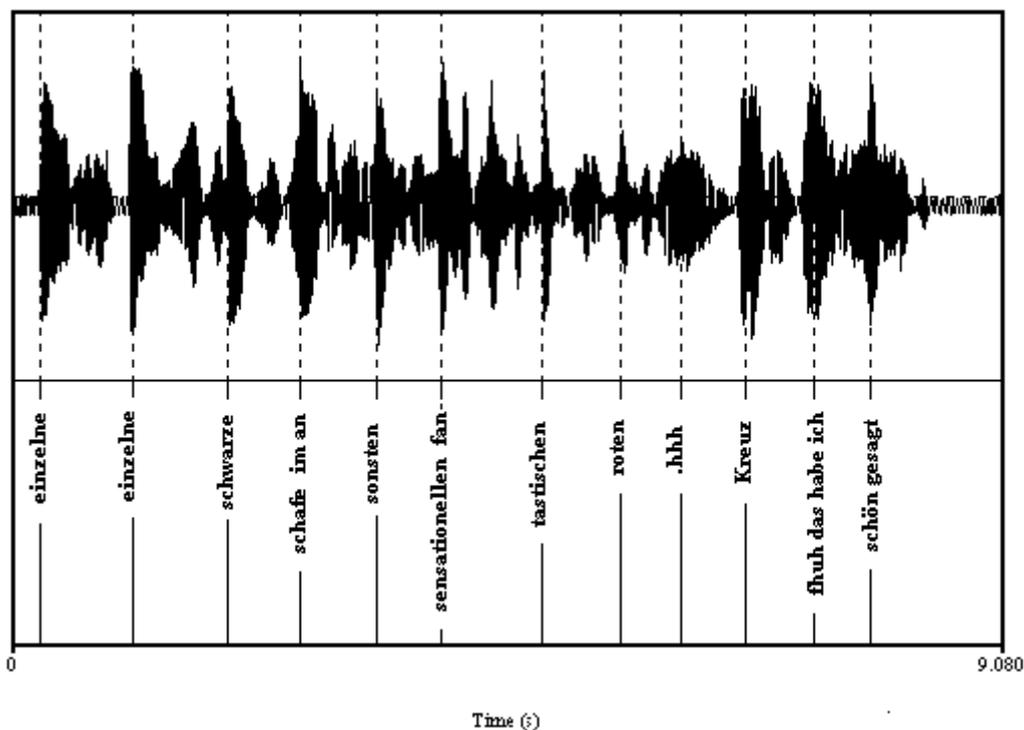
Creating a strong rhythmic pattern, Schmidt lines up the stressed syllables so they occur equidistant from each other. Like a beat in music the resulting isochronous rhythmic pattern can be tapped out with near metronome accuracy, as shown in the following diagram:

```

/einzelne /
/einzelne /
/schwarze /
/schafe im an- /
/sonsten /
/sensationellen fan- /
/tastischen /
/roten /
/.hhh /
/kreuz /
/fhuh das habe ich /
/schön gesagt

```

This auditory perception is additionally supported through the following wave form which shows that the points of greatest intensity line up at equal points in time:



Not only the rhythm, but also the pitch of the excerpt is highly patterned. Auditorily, one hears Schmidt starting the phrase in his upper voice range. With each stressed

syllable, the pitch falls in a patterned way from line 1 until Schmidt is in his lower voice range at line 10. At the end of the utterance, Schmidt is speaking in low volume (compared to the higher volume at the beginning of the utterance), at the bottom of the voice range, and, additionally, in a breathy voice². Coupled with the heavy in-breathing in line 9 and “fhuh” in line 11, the prosody contextualizes the extreme effort Schmidt is putting into his utterance in order to make it “schön”. Schmidt additionally makes use of alliteration, repeating the fricatives /s/ and /sch/ fifteen times in the brief excerpt.

How speakers form such stylized prosodic gestalts and the function these prosodic gestalts have in certain discourse activities is the object of research of the following dissertation. This two part goal is reflected in the title of the dissertation: “The Form and Function of Prosodic Stylization in Spoken Discourse”.

To deal with these two aspects the dissertation is divided into two sections: 1) a formal, descriptive section and 2) a functional section. The formal section intends to provide a conceptual and terminological apparatus for discussing stylized prosodic gestalts by applying principles of gestalt theory to prosodic gestalts. Complementing this section, the functional part of the dissertation intends to describe how and why speakers use stylized prosodic gestalts in certain activities within face-to-face interaction. Each of these two sections is again divided up into sub-sections, resulting in five chapters of the dissertation whose structure is as follows:

I. Formal, descriptive section

- A. Chapter 1: A Gestalt Approach to Prosody – Gestalt Principles and the Grouping of the Speech Continuum
- B. Chapter 2: The Interactional Relevance of Compound, Nested and Complex Prosodic Gestalts at, below, and above the Intonation Unit
- C. Chapter 3: A Gestalt Approach to the Highlighting and Stylization of Prosodic Structures

II. Functional section

- A. Chapter 4: The Placement and Stylization of *oh* in Informing Sequences
- B. Chapter 5: Prosodic Stylization in Assessment Sequences

² For a systematic description of voice quality see Laver (1980). Cf. also Crystal (1975).

Many current prosodic researchers either explicitly or implicitly deal with prosodic structures as perceptible auditory gestalts, including, among others, Auer (1996), Auer/Couper-Kuhlen (1994), Auer/Couper-Kuhlen/Müller (1999), Couper-Kuhlen (1990, 1993), Selting (1995), *et al.* Peter Auer in his article "Prosody and Syntax of Turn-Continuations" (1996) even asks directly: "Does prosody (in particular, intonation) build up gestalts independent from those in syntax?" (68). Auer's answer to this question is an undeniable "yes", but "if (...) the independence of prosody from syntax is considerable, the priority of syntax nonetheless cannot be denied either" (1996:75)³.

Although Auer does not specifically discuss gestalt theory or principles of gestalt theory, they do, in fact, underline his argument. Additional discourse material, which syntactically must be considered a second syntactic construction, may still be heard as a single prosodic gestalt. Such "tagged-on" material shows "prosodic integration" when one prosodic gestalt is heard and "prosodic exposure" when two gestalts are heard (Auer, 1996:70)⁴.

The application of organizing principles of gestalt theory to prosodic phenomena is more explicitly dealt with in Couper-Kuhlen (1990) and (1993). Couper-Kuhlen asserts that "perceptually isochronous sequences (...) have a number of similarities with auditory *gestalten*." Moreover, "the organizing principles of perceptually isochronous sequences are similar to those which Wertheimer proposed for *gestalten* in general" (1990:44). She then presents seven organizing principles proposed by Wertheimer (1923, 1938) and claims their relevance for rhythmic phenomena in speech. Couper-Kuhlen concludes that "there is good reason to think of perceptually isochronous sequences as auditory *gestalten*" (1990:47). Couper-Kuhlen ends the paper asserting a combined auditory-acoustic approach for the analysis of spoken English and stressing the importance of perception: "ultimately (...) the analyst's

³ Although I agree with Auer's assertion that syntax and prosody are independent of each other, I am reluctant to go as far as Auer in asserting the general priority of syntax. In at least one instance, Szczepek (2000) indicates that for signaling turn-completion, prosody seems to have priority over syntax.

⁴ See also Selting (1996).

decision is a perceptual one, which means that vague and ambiguous judgments must be reckoned with. But this is in the nature of rhythm as *gestalt*" (1990:53).

The auditory-acoustic approach was explored more rigorously by Couper-Kuhlen in her 1993 study of rhythm in speech. According to the study, even when acoustic measurements vary up to 20% between rhythmic intervals one may still perceive isochronous rhythm.

More recently Auer/Couper-Kuhlen/Müller's book Language in Time (1999) has discussed gestalt-psychological research on rhythm in spoken discourse. Supporting Couper-Kuhlen's 1990 findings, they assert that "gestalt psychology and subsequent research on perceptual thresholds for time and rhythm perception (...) suggest a cautious interpretation of instrumental measurements as the basis for an interactionally oriented interpretive approach to the meaning and function of rhythm. (...) what is called for is a combined auditory-perceptual and instrumental approach" (1999:14).

Such an auditory-acoustic approach, increasingly simplified through the availability of acoustic analysis software⁵, has helped advance prosodic research the last few years. Still, Auer states that "the available research on the role of intonation in conversation is scarce and the terminology undeveloped" (1996:68). I intend to show that the concepts and terminology of gestalt theory as proposed by gestalt theorists such as Max Wertheimer, Wolfgang Köhler, Kurt Koffka, Christian von Ehrenfels, *et al.* can benefit prosodic research. In the first half of the dissertation I wish to bring the two fields of prosody and gestalt psychology closer together, and, in a systematic way, show how prosodic structures are, in fact, governed by the same principles as other perceivable gestalts such as visual and musical gestalts. This formal, descriptive part intends to provide a useful apparatus for more precisely describing stylized prosodic phenomena.

⁵ The acoustic analysis software program "PRAAT, a system for doing phonetics on computer", was an indispensable tool for the following dissertation. For more information about this useful freeware program contact Paul Boersma at boersma@fon.hum.uva.nl

In chapter 1, “A Gestalt Approach to Prosody – Gestalt Principles and the Grouping of the Speech Continuum”, I argue for a gestalt approach for describing how speakers group the speech continuum into “compound”, “nested” and “complex” gestalts, thereby proposing an approach to the prosodic organization of discourse which describes a wider range of phenomena than current prosodic research allows. By applying the organizational gestalt principles of “proximity”, “similarity” and “directionality” (also called the principle of “smooth continuation”) in perceptual forms to prosodic phenomena, I will show that prosodic gestalts are often much more complex than the traditional “intonation contour” currently allows.

In chapter 2, “The Interactional Relevance of Compound, Nested and Complex Gestalts At, Below, and Above the Intonation Unit”, I wish to show that the intricate ways of grouping the speech continuum are not only a theoretical construct but are interactionally relevant for the interlocutors themselves. By displaying an orientation to groupings at, below and above the intonation unit, interactants themselves show the relevance of such grouping patterns in the interaction.

Continuing with a gestalt approach in chapter 3, “A Gestalt Approach to the Highlighting and Stylization of Prosodic Structures”, I wish to show the gestalt processes which allow speakers to manipulate the prosody of their spoken contributions into “highlighted” and “stylized” gestalts. A “highlighted” gestalt pattern describes a small part of a larger gestalt structure which has become salient via one or more of the gestalt processes below. A “stylized” gestalt describes the entire gestalt which has become salient. I assert three gestalt processes for forming such highlighted and stylized gestalts: 1) the “reduction” of irregular individual parts of the gestalt, 2) the “adjustment” of irregular individual parts into regular individual parts, and 3) the “ornamentation” of certain individual parts by overlaying additional gestalt elements on top of an underlying gestalt. Stylized prosodic gestalts are formed via a combination of these three processes on both local and global levels.

In the functional section of the dissertation I take a conversation analytic approach to show how stylized prosodic gestalts function as practices in certain kinds of discourse activities. In chapter 4, “The Placement and Stylization of *oh* in Informing Sequences,” I analyze the interactional effects that stylization and gestalt patterning

have on turn taking and floor holding; in chapter 5, “Prosodic Stylization in Assessment Sequences”, I consider the activity of assessing. In both of these activities it will be shown that prosodic stylization is relevant for the participants’ negotiation and interpretation of what is going on conversationally.

In chapter 4, “The Placement and Stylization of *oh* in Informing Sequences” I will describe how the placement of the *oh* in relation to the first turn informing and its prosodic stylization has specific interactional effects on turn taking and floor holding. I will show how placement and stylization are exploited by interactants as conversational practices in order to negotiate speaker’s rights. Moreover, I will show why these practices are specifically suitable in informing sequences to achieve particular interactional goals.

Chapter 5, “Prosodic Stylization in Assessment Sequences”, looks at the role of prosodic stylization within the activity of assessing, an activity which has been the focus of much work from a talk-in-interaction point of view. It will be shown that prosodic stylization may be exploited by interactants to express preference, even when, lexically, evaluative terms are equal. I will examine two kinds of assessment contexts in which prosodic stylization plays an important role in establishing rapport via co-participation in the activity and in allowing conversation participants to creatively respond to face-threats that occur in playful modalities.

With prosodic stylization interactants have a powerful interactional practice at their disposal. Speech can, and indeed often does, become musical, and conversation participants become like musicians. How and why this occurs will be explored in detail in the following.

2 Methodology and Theoretical Background

The dissertation takes an interactional perspective on the study of language. It combines the theoretical underpinnings of interpretive socio-linguistics and contextualization theory (Auer 1986; Auer/di Luzio 1992; Cook-Gumperz/Gumperz 1976; Gumperz 1982, 1992, 1992a) with the methods of conversation analysis

(Sacks, Jefferson, Schegloff 1974) and interactional prosodic research (Couper-Kuhlen, Selting 1996, Schegloff 1998).

2.1 Interpretive Socio-Linguistics and Contextualization Theory

From an interpretive socio-linguistic perspective, reality is seen as a social construct (Schütz 1932; Schütz/Luckmann 1975). The methods of contextualization theory are used in order to explain how meaning and reality are constructed by the conversation participants *in situ*. Naturally occurring interaction recorded on film or audio and enriched through ethnographic background information form the basis for the analysis.

John Gumperz, a leading defender of this perspective, asserts that “grammar and semantics cannot alone account for situated meaning” (1996:x). An additional aspect which must be considered are the indexical signs which point towards certain specific interpretations. Gumperz calls these indexical signs “contextualization cues” and defines them as follows:

A contextualization cue is one of a cluster of indexical signs (...) produced in the act of speaking that jointly index, that is invoke, a frame of interpretation for the rest of the linguistic content of the utterance. Such frames are subject to change as the interaction progresses and have different scopes, from the individual speech acts to sets of turns and responses, to entire social encounters.

(Gumperz, Levinson 1996:379)

In this quote it can be seen that contextualization theory is closely related to Goffman’s frame theory (1974). In both approaches interactants offer interpretive cues within the interaction in order to signal what is going on in the conversation.

With Gumperz’s 1982 groundbreaking work, Discourse Strategies, context became crucial for the understanding of how prosody functions. His chapter, “Prosody in Conversation”, attempts to show “how conversationalists use prosody to initiate and sustain verbal encounters” (1982:100). Indexical by nature, prosody is seen by

Gumperz as an important contextualization cue for guiding participants' understanding of the unfolding activity.

2.2 Conversation Analysis (CA)

A related interactive approach to the interpretive socio-linguistic perspective described above is conversation analysis as advocated by Sacks in his posthumously published lectures (1992), Sacks/Schegloff/Jefferson (1974, 1977), Sacks/Schegloff (1973) and Schegloff (1996) *et al.* One major difference between an interpretive socio-linguistic approach and a CA approach, however, is the inclusion of background ethnographic information. Whereas this is deemed necessary for ethnographic work, it is avoided by CA analysts. For CA researchers interested in the nuts and bolts of the organization of naturally occurring discourse at the micro-structural level, the inclusion of background information risks offering a pre-emptive interpretation of data⁶.

Influenced by Goffman's concept of "interaction order" (1964, 1967, 1983), CA analysts have attempted to create a method for uncovering and describing the practices and actions in the organization of spoken discourse. In this undertaking they have further developed the ethnomethodological perspective initiated by Garfinkel (1967). Exploiting the ability of audio and visual recording technology to capture social actions, thereby making them an observable research object, CA analysts have developed a method for uncovering the systematic organization of spoken interaction at the finest level. The creation of an inventory of social actions at the micro-level is seen as one of the main goals of CA methodology. This is expressed by Sacks in his lectures:

A base for using close looking at the world for theorizing about it is that from close looking at the world you can find things that we couldn't, by imagination, assert were there: One wouldn't know that they were typical, one might not know that they ever happened, and even if one supposed that they did one couldn't say it because an audience wouldn't believe it. Where, then, if we can add to the stock of things that can be theorized

⁶ For an introduction to CA methods cf. Deppermann 1999.

about we will have done something more or less important – if the things that we've added have any import to them.

(1992, 2:419-420)

2.3 Interactional Prosodic Research

In the so-called British tradition of prosodic analysis, researchers have attempted to uncover the communicative function of intonation. Such research, which frequently used constructed examples and which neglected the importance of context, often had the role of teaching intonation to second language users. This approach is best represented by Halliday (1967, 1970). Further attempts to uncover the communicative function of intonation can be seen in the work of Bolinger (1958), Quirk *et al.* (1972), Crystal & Davy (1975), Leech & Svartvik (1975), Van Ek (1975), Wilkins (1976), Munby (1978), Brazil, Coulthard & Johns (1980), and Tench (1990).

More recently, an interpretive socio-linguistic approach has been combined with conversation analytic methods and expanded to include its prosodic dimension in the interactional perspective proposed by Couper-Kuhlen/Selting (1996), Selting (1992, 1995)⁷. Whereas prosodic phenomena were previously dealt with only superficially at best in CA work, with the interactional approach proposed by Couper-Kuhlen/Selting it becomes central and is seen as an important aspect in the organization of discourse. Günthner describes the influence of CA on prosodic research:

Konversationsanalytische Arbeiten rückten erstmals systematische Untersuchungen von Phänomenen wie Zögerungspartikeln, Pausen, Abbrüchen, Lachen etc. in das Blickfeld ihrer Analysen. Prosodische Verfahren wurden hierbei allerdings nur am Rande berücksichtigt – und zwar speziell die Rolle der Intonation in Zusammenhang mit der Organisation des Redewechsels. Doch kam mit der Konversationsanalyse und deren systematischer Hinwendung zur Organisation alltagssprachlicher Interaktionen zugleich eine für die Interpretative Prosodieforschung zentrale methodologische Richtlinie auf: Zur Beschreibung und Erklärung prosodischer Verfahren ist es notwendig, diese im interaktionalen Kontext zu untersuchen und zu überprüfen, ob und wie sich die Interagierenden an diesen Parametern orientieren.

⁷ Cf. also Schegloff 1998.

(1998:47)

For the first time studies in conversation analysis began to consider the systematic analysis of phenomena such as hesitation particles, pauses, break offs, laughter, etc. Prosodic processes, however, were only marginally considered – most specifically the role of intonation in the organization of turn-taking. But CA and its systematic consideration of the organization of everyday speech interaction simultaneously offered central methodological guidelines for interpretive prosodic research: For the description and explanation of prosodic processes it is necessary to analyze them in the interactional context and to check whether and how the interactants orient towards these parameters.

(my translation)

With an interactive approach to prosody Couper-Kuhlen and Selting offer a way of describing a central and ever-present aspect of spoken interaction. Prosodic phenomena are indeed interactionally relevant, and the prosodic approach suggested by Couper-Kuhlen and Selting compensates for some of the deficits of other methodologies:

In proposing an interactional perspective on prosody as a ‘remedy’ for formal, functional and methodological problems in current research, we have taken a doubly empirical stance: (i) the approach advocated takes empirical data as its object of study, and (ii) it seeks empirical evidence for the validation of its analyses. While the analysis of empirical data is arguably not new in prosodic studies, the kind of empirical validation suggested here is often absent from other approaches.

(1996:47)

3 The Data

The claims which I make in this dissertation are empirically grounded; i.e., the phenomena were discovered in the data after repeated and careful listening. The empirical basis of the analyses consists of 19 recordings of approx. 9 hours of recorded data. These recordings were digitalized as *wav* files in order to be analyzed on the computer. The recordings consist of informal conversations among friends, lovers, and family members, as well as recordings of radio talk show programs in

which a moderator speaks with invited guests in the recording studio as well as with listeners who call in. In a few cases, the moderator speaks directly to the radio audience. Among the informal conversations, much of the data consists of telephone conversations. Other conversations were recorded at informal gatherings of friends and family. Following is a list of the recordings:

- 1) Radio and television programs:
 - a) Barbara Carlson (BC) on Death 0:46:44
 - b) BC with Congressman Graham 0:16:41
 - c) BC with the Dog-lady 0:10:06
 - d) BC in the Hot-tub 0:17:41
 - e) BC Saving the Fat and Ugly 0:13:10
 - f) The Pickle Queen 0:14:10
 - g) Paul Harvey 0:16:36
 - h) Rush Limbaugh 0:29:09
- 2) Telephone conversations
 - a) Cutie Pie 0:25:53
 - b) Happy Anniversary 0:03:57
 - c) Moped 0:01:54
 - d) Party 0:04:55
 - e) Power Tools 0:08:44
 - f) Newport Beach (28 separate telephone conversations) 3:20:00
- 3) Informal Conversations
 - a) Cuz 0:27:00
 - b) Appease the Monster 0:27:56
- 4) Miscellaneous
 - a) God's Eternal Home 0:01:46
 - b) Preacherly Vine 0:01:47

Many of the recordings above are from the University of California, Santa Barbara corpus of spoken English. I am very thankful for having been given access to this data. The other recordings are from the corpus of spoken data at the Universität Konstanz as well as recordings from my own private collection.

The recordings from Santa Barbara were originally transcribed according to the conventions proposed by DuBois *et al* (1993). Relevant transcripts from these recordings have been rewritten according to GAT (Gesprächsanalytisches Transkriptionssystem) conventions as proposed by Selting *et al* (1998). In some cases, the original transcriptions were retained or slightly changed according to my own auditory perception of the data. For an explanation of the transcription conventions see “Appendix: Basic Transcription Conventions”, p. 203.

Chapter 1: A Gestalt Approach to Prosody – Gestalt Principles and the Grouping of the Speech Continuum

The following argues for a gestalt approach for describing how speakers in everyday spoken discourse form their contributions into “compound”, “nested”, and “complex” structures. From a gestalt perspective, the grouping of auditory events is seen to be a more intricate phenomenon than current prosodic research provides for. The traditional concept of the “intonation unit” or “intonation phrase” – as described by such prosodic linguists as Crystal (1968), Cruttenden (1986), Couper-Kuhlen (1986) Selting (1995) *et al.* and as used by most researchers analyzing prosodic phenomena – is but one way of grouping the speech continuum among others found in real conversation. By applying the organizational gestalt principles of “proximity,” “similarity,” and “directionality” in perceptual forms – as outlined by gestalt psychologist Max Wertheimer (1923, 1938) for visual gestalts and Fred Lerdahl and Ray Jackendoff (1983) for musical gestalts – to auditory prosodic gestalts, I propose an approach which more precisely describes a wider range of phenomena than currently considered in prosodic research.

1 Introduction

Christian von Ehrenfels belonged to the first generation of psychologists in the late 19th century. His interest in perception led him to propose the concept of gestalt in his seminal essay, “On Gestalt Qualities” (1890). Von Ehrenfels’ work has since had a major impact on a number of fields, including psychology, philosophy, linguistics⁸, physics, sociology, music, etc. According to von Ehrenfels, “by a gestalt quality⁹, we understand a positive content of presentation bound up in consciousness with the presence of complexes of mutually separable (i.e. independently presentable)

⁸ Many linguists – such as René Thom, Jean Petitot, Wolfgang Wildgen, Ray Jackendoff *et al.* – have dealt with “semantic” rather than prosodic gestalt structures. Although such structures have some things in common with visual and auditory gestalts, such an approach weakens the notion of gestalt as perceptible structures. For an overview of such a gestalt linguistic perspective, see Smith (1988), Petitot (1985, 1986) and Wildgen (1985). Linguists from an interactive perspective have begun to consider prosodic structures as gestalts. See, for example, Auer (1996), Auer/Couper-Kuhlen (1994), Couper-Kuhlen (1990, 1993), Selting (1995) *et al.*

⁹ Instead of von Ehrenfels’ term “gestalt quality”, most psychologists today simply use “gestalt”. See Kurt Grelling/Paul Oppenheim (1937/1938).

elements. That complex of presentations which is necessary for the existence of a given gestalt quality we call the foundation (*Grundlage*) of that quality” (1890:93).

With this paper von Ehrenfels, influenced by Ernst Mach’s observations (1886), established the basis of gestalt psychology which was further developed in the 1920s and 1930s by the “Berlin School of Gestalt Psychology”, principally represented by psychologists such as Max Wertheimer (1922, 1923, 1925, 1925a, 1927, 1933), Wolfgang Köhler (1920, 1923, 1924, 1925, 1925a, 1929, 1929a, 1930), and Kurt Koffka (1922, 1923, 1924, 1924/1925, 1928, 1931, 1935)¹⁰. Although von Ehrenfels’ original essay focused mainly on temporal musical gestalts, the psychologists of the Berlin school focused primarily (although not exclusively) on non-temporal visual gestalts. In more recent research, Fred Lerdahl and Ray Jackendoff (1983) have applied and further developed the ideas of the gestalt psychologists to musical structures.

The prosodic elements of speech – pitch, duration, and volume – are gestalt elements. As such they are the same gestalt elements found in music and subject to the same principles of perception as musical elements. Like musical elements, they are responsible for grouping phenomenal events. Lerdahl and Jackendoff’s definition of the term “group” within musical gestalts can equally be applied to prosodic structures. Groups are organizational units of sound signals, “such as motives, themes, phrases, periods, theme-groups, sections, and the piece itself” (1983:12). In prosody, linguists have described interactants’ division of the speech continuum in terms of the intonation unit. However, as will be shown, prosodic groups may be much more intricate, and, as in music, there can be groups within groups, which are grouped within other groups; there may be parts of groups that overlap parts of other groups, etc.

Following a discussion of the traditional unit of spoken discourse, the intonation group, as defined by such researchers as Alan Cruttenden (1986), David Crystal (1969) and presented by Couper-Kuhlen (1986), I will propose a gestalt approach to prosody by applying the gestalt principles of proximity, similarity and directionality to

¹⁰ A helpful online resource containing many of the most influential gestalt texts is run by the Society for Gestalt Theory and its Applications at <http://www.enabling.gestalt/gerhards/gtax.html>.

actual spoken discourse. Afterwards, I will show how these principles are exploited to form auditory patterns within the speech continuum that are more complex than currently allowed by traditional prosodic research with basic 'intonation unit'.¹¹

1.1 The Intonation Unit

The typical group of spoken discourse has been variously termed and characterized by a number of researchers. It has been called the "breath-group", the "sense-group", the "tone-group", the "tone-unit", the "phonological clause", the "idea unit", the "information unit", and, more recently, the "intonation group", the "intonation phrase" and the "intonation unit". The concept of the intonation unit has been taken over and applied by researchers from an interactional perspective to empirical data and proves an important and powerful linguistic construct for describing the organization of the speech continuum.

Alan Cruttenden offers six criteria for establishing the boundaries of intonation groups. The first four are external criteria; the last two are internal: 1) pause, 2) quick anacrustic syllables denoting the beginning of a new intonation group, 3) lengthening on the final syllable of the intonation group, 4) a change in pitch direction of an unaccented syllable, 5) the existence of at least one stressed syllable, 6) a pitch movement to or from an accented syllable (1986). According to Cruttenden, "one or both of the following criteria will in most cases delimit intonation-groups: (i) change of pitch level or pitch direction of unaccented syllables (ii) pause, and/or anacrusis, and/or final syllable lengthening, plus the presence of a pitch accent in each part-utterance thus created" (1986:42).

Couper-Kuhlen offers a model of the internal structure of the tone-unit which is based on Crystal (1968:208):

¹¹ In the following, the terms 'intonation unit' and 'intonation group' will be used interchangeably.

		nucleus	
	head	nucleus	
	head	nucleus	tail
pre-head	head	nucleus	
pre-head		nucleus	
pre-head		nucleus	tail
		nucleus	tail
pre-head	head	nucleus	tail

(Couper-Kuhlen, 1986:78)

Ashby schematizes the above as follows:

tone-unit → (pre-head) (head) nucleus (tail)

(quoted in Couper-Kuhlen, 1986:78)

In this diagram, eight possible variations for the internal structure of a tone-unit are proposed. Obligatory for all variations is the nucleus, defined as “the most prominent syllable in a tone-unit” (1986:79). This may be preceded by the head, which consists of all syllables starting from the first stressed syllable before the nucleus up to the nucleus. The head may in turn be preceded by the “pre-head”, which consists of all unstressed syllables preceding the first stressed syllable of the head. The nucleus may then be followed by the tail, defined as “any stressed and/or unstressed syllables following the nucleus”.

One problem with the intonation unit is that in actual spoken discourse it often proves quite difficult to identify unambiguously – a problem attested to by many a researcher transcribing spoken data. Brown, Currie and Kenworthy (1980) assert this problem in their book Questions of Intonation: “(...) we encounter constant difficulty in identifying tone groups in spontaneous speech (...)” (in Cruttenden 1986).

In spoken discourse speakers often speak in intonation units as described above, but not exclusively. The units heard in actual spoken discourse are much more varied than the eight variations proposed above. Deviations from strict intonation units have frequently been considered incomplete or ill formed. Cruttenden, for example, asserts that the inherent difficulty in identifying clear-cut intonation group boundaries in spoken discourse is due to the “broken nature of much spontaneous speech, including as it does hesitation, repetitions, false starts, incomplete sentences and sentences involving a grammatical caesura in their middle” (1986:35). From a conversation analytic perspective, however, where discourse is thought to be finely ordered at all levels, Cruttenden’s explanation is far from satisfactory.

From the gestalt perspective which I propose, the intonation unit is one particular kind of auditory gestalt in which the beginning and end boundaries of a sequence of phenomenal events are clearly and congruently marked. Cruttenden’s description of boundary markers as well as Crystal’s model of the internal structure of the intonation unit may well be applied to such gestalts. There are other kinds of gestalts, however, which exist alongside the intonation unit. Such structures may be smaller or larger than the intonation unit. They may also have a more complicated organizational structure than the traditional intonation unit. From the gestalt perspective proposed here, such gestalts, often described as ill formed or incomplete by researchers insisting that every spoken utterance fit the criteria of an intonation unit, do, in fact, exist and are exploited and oriented to by speakers.

1.2 A Gestalt Approach to the Prosodic Grouping of the Speech Continuum

By applying the gestalt principles of proximity, similarity and directionality to prosodic gestalts, it will be seen that there are auditory groups above, at and below the intonation phrase. Furthermore, these groups will be seen to have organizational structures which are often more complex than that of the intonation unit.

The gestalt principles of proximity, similarity, and directionality dictate how human beings perceive the grouping of phenomenal elements (Wertheimer 1923, 1938). These three principles, shown to be the guiding principles behind group formation in

visual and musical gestalts, are also responsible for group formation in prosodic gestalts.

To illustrate the gestalt principles I will first present visual examples taken from gestalt psychologist Max Wertheimer in his groundbreaking essay “Laws of Organization in Perceptual Forms”¹². I will then present musical examples taken from Jackendoff and Lerdahl (1983), who apply Wertheimer’s principles to musical gestalts. I will then show how these three gestalt principles are exploited in the prosodic grouping of the speech continuum in an identical way to Wertheimer’s visual and Lerdahl and Jackendoff’s musical gestalts.

2 The Proximity Principle

According to gestalt psychologists, proximity is the most important principle involved in the delimitation of groups (Wertheimer, 1923:77; Lerdahl/Jackendoff 1983:41-42). The principle says that individual objects that are in close proximity to one another will be perceived as a group; objects that are further apart will not. For visual gestalt phenomena proximity refers to the spatial distance between two objects; for events which unfold in time – such as music or spoken language – proximity refers to the amount of temporal distance between two auditory events. Spoken language, like music, unfolds in time; therefore, filled and unfilled pauses become a central aspect of the grouping of the speech continuum.

The following examples demonstrate the principle of proximity in visual gestalts (2.1.1, 2.1.2) and musical gestalts (2.2.1, 2.2.2). Afterwards, examples 2.3.1 and 2.3.2 will demonstrate how prosodic gestalts in spoken discourse are similarly subject to the principle of proximity.

¹² First published as “Untersuchungen zur Lehre von der Gestalt II”, in *Psychologische Forschung*, 4, 301-350 (1923). Translation published in Ellis, W. (1938) *A Source Book of Gestalt Psychology* (71-88). London: Routledge & Kegan Paul.

2.1 Proximity Principle in Visual Gestalts

Consider the following examples of visual gestalts, based on the findings of Wertheimer (1923, 1938) and presented by Lerdahl and Jackendoff (1983):

2.1.1 Left Grouping Circles



2.1.2 Right Grouping Circles



In 2.1.1 the principle of proximity causes the viewer to perceive the first and second circles as one group, and the third circle – separated by a greater distance from the second circle than the first circle – as another. Similarly in 2.1.2, the second and third circles are grouped together because of the shorter spatial distance between them. The first circle, which is further away, is grouped apart. In both examples, the principle of proximity delimits those objects which are closer together as a group. Those objects which are further apart are perceived as their own group.

2.2 Proximity Principle in Musical Gestalts

Lerdahl and Jackendoff (1983) have found the same principle of proximity operating in musical gestalts. Unlike in visual gestalts, musical phenomenal events are temporal. As stated above, proximity in temporal gestalts refers to the distance in time that occurs between auditory phenomena, in this case, musical notes. Consider the following musical gestalts:

2.2.1 Left Grouping Quarter Notes



2.2.2 Right Grouping Quarter Notes



(Lerdahl and Jackendoff 1983:39-40)

Again, one sees the principle of proximity operating. As with the visual gestalt of 2.1.1, a listener of 2.2.1 would perceive the first two quarter notes as a group, the third quarter note following the quarter pause would be perceived as a separate group. The same holds true for 2.2.2, where the second and third quarter notes following the quarter pause would be perceived as a group and the first quarter note would be perceived as its own group.

2.3 Proximity Principle in Prosodic Gestalts

Prosodic gestalts, like musical ones, are composed of auditory events which unfold in time. As with musical gestalts, auditory events within close proximity of each other are heard as belonging together; those separated by a pause are grouped apart. Consider the following excerpt:

2.3.1 “Good Morning”

Radio talk-show moderator Barbara Carlson (B) in her radio program “Barbara Carlson and Friends” answers a call from a phone-in caller (C):

B: Mike good morning
C: (.) uh yes (.) uh

Although short, this excerpt, like the visual gestalt in 2.1.1 and the musical gestalt in 2.2.1, nicely shows how the principle of proximity works to delimit gestalt boundaries. In line 2, C utters “uh” and “yes” one after the other in close proximity.¹³ “Yes” is then followed by a pause, which is followed by the second token, “uh”. One clearly perceives the first two elements, “uh” and “yes” as a group. The third element, “uh”, is perceived as its own group.

To complete the picture, here is a correlate example to visual gestalt 2.1.2 and musical gestalt 2.2.2, in which the final two discourse elements are grouped together, and the first element is perceived as separate:

2.3.2 “Wow”

Radio moderator Barbara Carlson (B), in her talk show program “Barbara Carlson and Friends”, discusses assisted death.

B: that you anticipate what’s going to happen
and then it just hits you
(.)
and you have to do it A:LL OVer aGAIN
.hhh (.) wow hh.

In line 4 one finds three individual phenomenal events: audible in-breathing, the interjection “wow”, and audible out-breathing. Following the heavy breathing is a pause, which is followed by the interjection “wow”. Audible out-breathing immediately

¹³ I am assuming that the basic prosodic elements are syllable-like in nature.

follows this¹⁴. The principle of proximity dictates that the interjection and the out-breathing are perceived as a group. The in-breathing is perceived as its own group.

As shown above, the prosodic element of duration functions to group speech phenomena according to the gestalt principle of proximity. In the following, I will add a correlate to this principle and demonstrate the correlate with visual, musical and prosodic examples.

3 The Proximity Correlate

According to gestalt principles, the proximity principle groups and separates individual phenomenal events according to the spatial or temporal distance between them. A correlate to this principle asserts that the closer the proximity, the stronger the perception that the elements belong together; the wider the proximity, the stronger the perception that the individual elements are separate. Lerdahl and Jackendoff (1983) describe the proximity correlate as follows:

The grouping effect can be enhanced by exaggerating the difference of distances. It can be weakened by reducing the disparity.

(Lerdahl and Jackendoff 1983:40)

I will now demonstrate the proximity correlate using visual examples from Wertheimer, musical examples from Lerdahl and Jackendoff, and prosodic examples taken from my corpus of data of everyday spoken discourse.

¹⁴ Traditional prosodic research would not include audible breathing as part of the intonation unit. Many scholars would consider breathing, along with coughing, sneezing, hiccupping, etc. as uncontrollable and therefore paralinguistic. Undeniably, however, these phenomena are audibly perceptible and therefore subject to gestalt principles. The “uncontrollable” nature of such phenomena is also questionable, as demonstrated by anyone who has ever politely coughed in order to get a waiter’s attention.

3.1 Proximity Correlate and Visual Gestalts

Consider the following visual examples, which are based on Wertheimer (1923, 1938) and reprinted in Lerdahl and Jackendoff (1983):

3.1.1 Left Grouping Circles



3.1.2 Weakened Perception of Left Grouping Circles



3.1.3 No Grouping of Circles



(Lerdahl and Jackendoff 1983:40)

In the above examples the relative distance between the middle circle and the outer two circles changes. In 3.1.1 the principle of proximity is clearly at play; consequently, there is a strong perception that the first two circles form a group and that the final circle, separated from the others by a larger distance, is on its own. In 3.1.2 the perception correlate begins to influence perception: the first two circles are not as strongly grouped as in 3.1.1. Still, one has a sense that the first two circles belong together. This changes, however, in 3.1.3, in which one can no longer group two of the three circles. The three circles, which are at an equal distance from each other, no longer display the same grouping structure as 3.1.1 or, to a lesser extent, 3.1.2.

3.2 Proximity Correlate in Musical Gestalts

Lerdahl and Jackendoff (1983) have shown that the proximity correlate also has an important function in the perception of temporal musical gestalts, as the following examples demonstrate:

3.2.1 Left Grouping Eighth Notes



3.2.2 Weakened Perception of Left Grouping Eighth Notes



3.2.3 No Grouping of Notes



(Lerdahl and Jackendoff 1983:40)

Just as the grouping structure of the visual gestalts in 3.1.1 – 3.1.3, the grouping structure of the above musical gestalts weakens in 3.2.2 and dissolves completely in 3.2.3. In 3.2.1 an eighth note is followed by an eighth note, which is followed by a quarter pause, which is followed by a quarter pause, which is followed by a quarter note. One perceives the first two eighth notes as a group; the final note is clearly heard as separate.

In 3.2.2 the strong grouping perception begins to break down. An eighth note is followed by two eighth pauses, which are followed by an eighth note, which is followed by three eighth pauses, which are followed by an eighth note. The notes are not yet equally spaced; therefore, as the time increases between the first and second

note and decreases between the second and third notes, one still tends to group the first two notes together and the third apart.

The proximity correlate predicts that no sense of grouping occurs when the notes are equally distant from one another. This is demonstrated in 3.2.3 in which a quarter note is followed by an eighth pause, which is followed by a quarter note, which is followed by an eighth pause, which is followed by a quarter note.

In the following, the proximity correlate will be seen to play a role in the perception of prosodic groups.

3.3 Proximity Correlate in Prosodic Gestalts

The following examples show in the same way as the above visual and musical examples how auditory phenomenal events in spoken discourse are subject to the proximity principle correlate. Consider the following excerpt:

3.3.1 “KSTP”

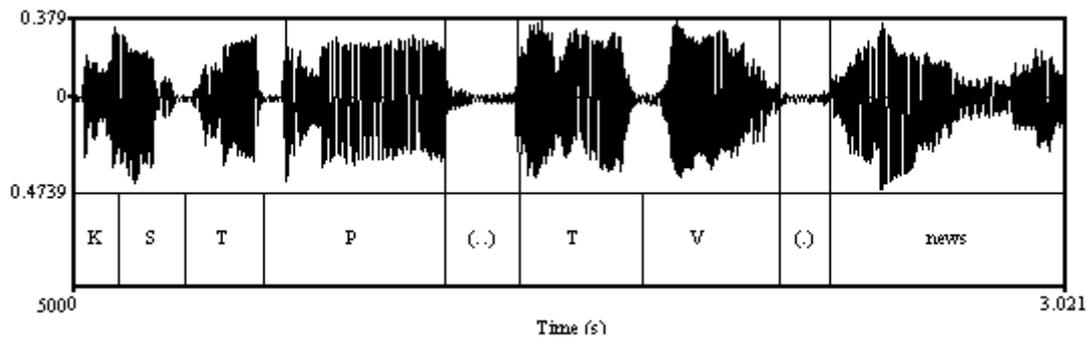
Radio moderator Barbara Carlson (B) has just taken a call from “Jeff”.

```
J: I was watching  
→ KSTP (...) TV (.) news  
B: yes
```

Line two consists of seven syllables: “k”, “s”, “t”, “p”, “t”, “v”, and “news”. The first six syllables are letters, the last syllable is a word. Due to the proximity principle, one distinctly hears three groups:

- 1) “k”, “s”, “t”, “p”
- 2) “t”, “v”
- 3) “news”.

Between group one and group two there is a longer pause than between group two and group three. This can be seen in the following wave form:



Due to the proximity principle the first group is perceived as being a group clearly distinguished from the second group. The beginning of the syllable “p”, which is held and then followed by a pause, is further away from the beginning of the next syllable “t” than it is from the first syllable of the first group, “k”. It can be argued that in order to distinguish the boundary between the letters of the radio station “KSTP” from the letters of the acronym “TV” the speaker makes use of a longish pause.

Between the second and third groups there is a shorter pause than between the first and second groups, and, in fact, one hears these two groups as being closer together than the first and second groups. The proximity correlate allows for a stronger perception of group separation between “KSTP” and “TV” than between “TV” and “news”. I argue that this longer pause is used to avoid the potential confusion resulting from the grouping of the first six syllables (which are also letters) together.

The proximity correlate influences the perception of patterns not only on the syllable level but also on larger gestalts that consist of a number of syllables. This holds true for all of the perception principles. Consider the following example:

3.3.1 “John Derris”

Radio moderator Barbara Carlson (B) takes a phone call from the next caller, who turns out to be Minnesota State Senator John Derris (D).

1 B: John good morning you're on KSTP
 2 (...)
 3 J: good morning
 4 B: hello sir
 5 (...)
 6 J: hel[lo
 7 B: [welcome to the hot-tub
 8 (.)
 9 J: hello
 10 B: hello
 11 J: this John Derris
 12 B: oh John Derris

This example contains long and short pauses which break up the excerpt into four gestalts, some of which are more strongly perceived as individual gestalts than others due to the proximity correlate. The example can therefore be seen as a prosodic correlate to visual gestalts 3.1.1 and 3.1.2 as well as to musical gestalts 3.2.1 and 3.2.2 above.

Following line 1, in which Barbara Carlson greets her phone-in caller for the first time, is a very long pause, after which John Derris offers his opening in line 3. His greeting is immediately followed by Carlson's second attempt at a greeting, which is followed by yet another very long pause. Because of the long pause in line 2, one clearly hears Carlson's "John good morning you're on KSTP" as separate from Derris' next turn in line 2 "good morning" and Carlson's turn in line 4 "hello sir". Furthermore, the long pause in line 5 also strongly separates it from what follows.

Following the long pause in line 5, Derris again makes a second attempt at a greeting. This is followed in partial overlap by Carlson's greeting, "welcome to the hot-tub"¹⁵, which is followed by a shorter pause, which is then followed by John Derris' third attempt at a greeting in line 9. Because of the long pause in line 5, one has the perception that lines 6 and 7 are clearly separate from lines 3 and 4. Because of the shorter pause in line 8, however, one has a weaker perception that the lines are their own group. Line 9 is immediately followed without a pause by B's fourth attempt at a greeting, which is followed without pause by John Derris'

¹⁵ In this particular episode of "Barbara Carlson and Friends", Barbara Carlson is moderating her program while sitting in a hot-tub at the Minnesota State Fair.

introduction of himself, which is followed without pause by Barbara Carlson's recognition of who the caller is in line 12.

Barbara Carlson begins the activity of greeting in line 1¹⁶. Instead of the preferred second to this adjacency pair, there is a long pause before Derris offers his greeting. This is immediately followed by Carlson's second greeting. Due to the proximity correlate, Derris' "good morning" is not unequivocally heard as the second turn to Carlson's first turn in line 1. There is simply too much time between the two greetings. Derris' "good morning" may be heard (and is indeed reinterpreted by Carlson) as the first turn greeting in a new adjacency pair. Carlson completes the new greeting activity in line 4 with her second turn "hello sir". The confusion, however, is still not cleared up, and in lines 6 and 7 Carlson and Derris again offer greeting tokens, this time in overlap. Due to the overlap, it is still unclear whose turn was the first of the adjacency pair and whose was the second. Following another long pause, Derris offers yet another first turn greeting in line 9: "hello". This is unmistakably heard by Carlson as the first turn, as she offers her second turn, "hello", rhythmically integrated. Having finally cleared up the confusion, the conversation is able to continue past the greeting.

One can draw a visual correlate of the prosodic grouping in 3.3.1 as follows, where each intonation unit is represented by a circle:

3.3.2

○ ○○ ○○ ○○○○

To summarize, the proximity correlate asserts that the greater the distance between two phenomenal events, the stronger one perceives that the events are their own groups. The shorter the distance, the stronger the perception that they belong together. In the example above, one has the impression of four gestalts. The pauses between the first and second, and the second and third gestalts, however, are much greater than the pause between the third and fourth gestalt. One perceives the third

¹⁶ For a CA description of conversational openings in radio talk show programs cf. Panese 1992.

and fourth gestalts almost as a single group. The first and second gestalts are clearly their own groups.

4 The Similarity Principle

According to the principle of similarity in gestalt theory, those items which are similar are grouped together; those items which are dissimilar are grouped apart. The following examples will illustrate the principle of similarity in visual, musical, and prosodic gestalts.

4.1 Similarity Principle in Visual Gestalts

The following two examples of visual gestalts each contain five shapes which occur at an equal distance from one another; therefore, the principle of proximity does not contribute to the delimitation of group boundaries. However, one still has a strong perception that there are two groups:

4.1.1

□□□○○

4.1.2

□□○○○

(Lerdahl and Jackendoff 1983:40)

In 4.1.1 the three squares to the left are perceived as a group, and the two circles to the right are perceived as a group. In 4.1.2 the two squares to the left are perceived as a group, and the three circles to the right are perceived as a group. The principle of similarity delimits the boundaries via likeness.

4.2 Similarity Principle in Musical Gestalts

Just as similar visual shapes are grouped together when they are equally distant from one another, so are musical notes when they have the same pitch, and when they are temporally spaced equally apart. Lerdahl and Jackendoff offer a musical equivalent to the above visual gestalts.

Consider the following two examples 4.2.1 and 4.2.2:

4.2.1



4.2.2



(Lerdahl and Jackendoff 1983:40)

In both examples all the quarter notes occur at an equal distance from one another. The principle of similarity, however, makes a group of the first three quarter notes in 4.2.1, which have lower pitch than the second two quarter notes (and indeed the same low pitch). Similarly, in 4.2.2 the first two notes are higher than the succeeding three quarter notes. Although all notes are equidistant from one another, one perceives the higher notes as a group and the lower notes as a group.

4.3 Similarity Principle in Prosodic Gestalts

The above musical examples (4.2.1 and 4.2.2) show how similar pitches of various individual notes will group those notes together according to the similarity principle.

This same phenomenon is not limited to music, but also functions in prosodic gestalts. This shall first be demonstrated on the syllable level in the following example 4.3.1, and then again on the level of the intonation unit in example 4.3.2:

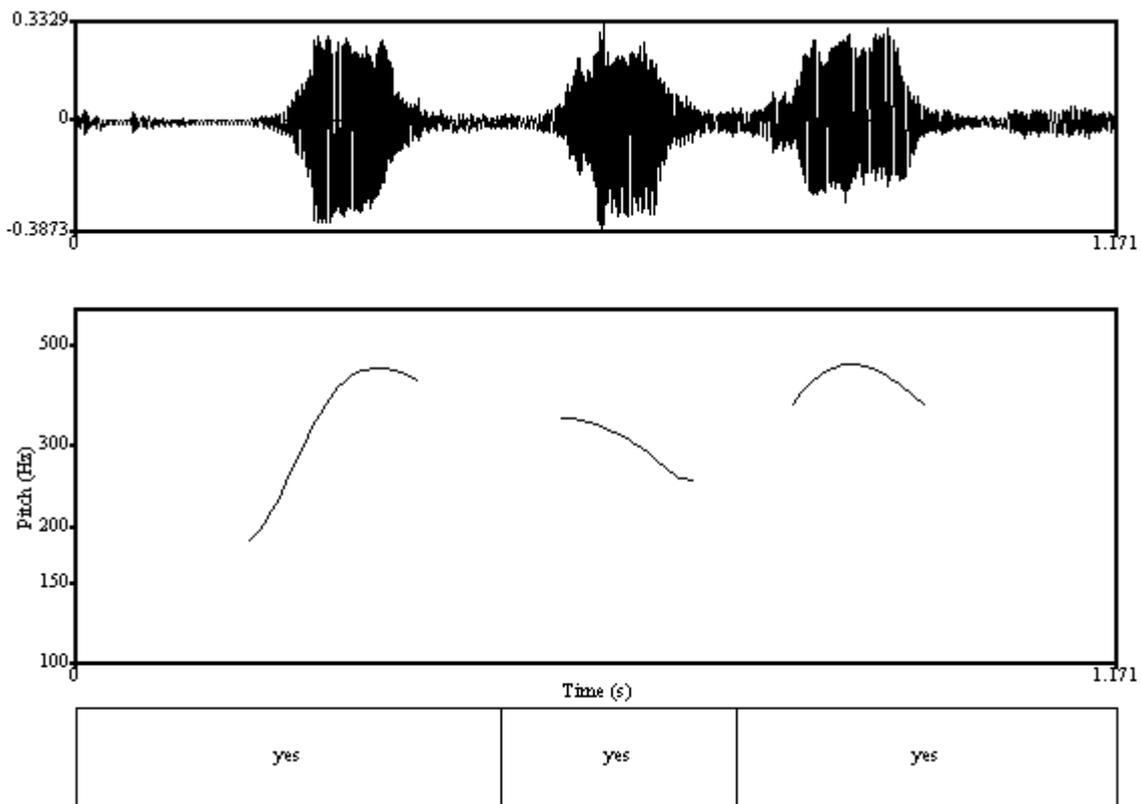
4.3.1 “Yes yes yes”

In this excerpt from the Barbara Carlson and Friends Show, which is being broadcast from the Minnesota State Fair, the moderators Peter (P) and Barbara (B) discuss the increasing age of the speaker of the House, Irv Anderson.

1 B: because those democrats
I'm I'm asking you
P: you know [he could
B [that's the
5 P: [Irv Anderson could get in here
B: [question you know we can put him in ??
P: free today
because it's seniors' [day
B: [yes
10 and he could [go see Lawrence Welk
B: [seventy years old
P: he's [that old
B: [yes
→ yes yes yes
15 .hh but you see he was a down and dirty fighter

In this excerpt both speakers frequently speak loudly and simultaneously as they battle for the floor. P and B's utterances overlap in lines 3 and 4, lines 5 and 6, lines 8 and 9, lines 10 and 11, and lines 12 and 13. In line 9 and in line 13 B shows agreement with P with her contribution “yes”, which, in both cases, occurs in overlap with the end of P's intonation phrase. In line 14 B says “yes” three times. This line, and the following lines, display no overlap. B has secured the floor for herself. In the following line she audibly inhales and then immediately shifts her prosody, which is lower in pitch, less loud, and slower.

Line 14, which secures B the floor, consists of three syllables: “yes”, “yes”, and “yes”. The first and last syllables are realized with the same pitch: approx. 480 Hertz. The pitch of the second “yes” is lower: approx. 350 Hertz. This can be seen in the following acoustic analysis:



Due to the principle of similarity, one hears the third “yes”, which is realized at the same pitch as the first “yes”, as belonging to the first “yes”.

Like the principle of proximity as described above, the principle of similarity can influence the perception of phenomenal events at various levels of complexity. Not only can one observe this principle at the syllable level (cf. 4.1.3), one can also see it at the phrase level. Consider, for example, the following:

4.3.2 “We’re All Gonna Die”

Radio moderator Barbara Carlson (B), in a phone-in program concerning death and dying, talks to her radio audience.

- 1 B: we’ve all been born
- 2 (.)
- 3 <<1, pp> and we’re all gonna die>
- 4 (.)
- 5 have you thought about it

Whereas lines 1 and 5 are spoken in a normal pitch register, the register of line 3 suddenly falls. The lower pitch register of line 3 sets the intonation phrase of this line apart from the preceding one in line 1. Similarly, the higher pitch register of line 5 in comparison to line 3 separates this intonation unit from the one in line 3.

In line 5 Carlson returns to the prosody she uses in line 1, and due to the principle of similarity, one hears this intonation unit as connected to the intonation unit in line 1. It can be argued that Carlson returns to the same lighter modality of line 1 after a short, very serious modality in line 3. Prosody seems to be a frequently used practice for signaling such quick shifts in modality.

A visual equivalent of example 4.3.1 and 4.3.2 would look as follows:



And a musical equivalent of example 4.3.1 and 4.3.2 would sound as follows:



Some may argue that the pauses in lines 2 and 4 are what separate line 3 from the surrounding discourse, and, in fact, the pauses do group the three lines apart. However, this does not explain the full complexity of the grouping structure. Without the low pitch register, which sets off line three from lines 1 and 5, the grouping structure would simply be (x) (x) (x). A visual equivalent would be ○ ○ ○. As line 3 is lower in pitch, however, it is additionally grouped off within this grouping structure: (x) [(x)] (x). Not only is line 3 grouped off by pauses, but also the lower pitch register puts it on a different level of organization than lines 1 and 5.

5 The Directionality Principle

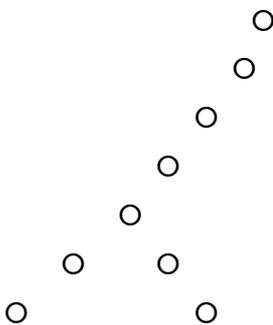
The principle of directionality is the third gestalt principle discussed by Wertheimer (1923, 1938). After presenting a few examples of visual and musical gestalts, I will present some examples of prosodic gestalts, in which the principle of directionality governs the delimitation of the boundaries of prosodic groups.

5.1 Directionality Principle in Visual Gestalts

According to the principle of directionality (also referred to as the “principle of smooth continuation”), phenomenal units are grouped together when the individual parts appear to line up together and continue smoothly in a single direction. In other words, when the individual phenomenal events align so as to form a pattern in which the whole gestalt implies a certain direction, the observer perceives the continuation of the contour. One therefore perceives the elements as a group.

For an example of the principle of directionality in visual gestalts, consider the following¹⁷:

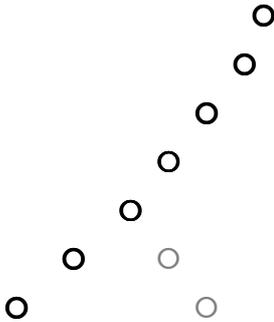
5.1.1



¹⁷Example adapted from "ABC Graphic Design: Gestalt Psychology."
<http://www.fa.indiana.edu/~mkato/document/graphic>

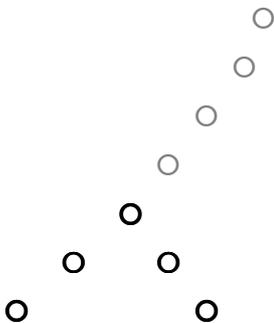
According to the principle of directionality, one perceives the seven circles which form a single contour from bottom left to top right as a group, as demonstrated below in 5.1.2:

5.1.2



The principle of directionality prevents, however, a grouping such as in 5.1.3:

5.1.3



The delimitation does not occur according to the principles of proximity or similarity but according to the principle of directionality or "good continuation".

In the following, I will demonstrate how the principle of directionality also occurs within musical gestalts.

5.2 Directionality Principle Musical Gestalts

The principle of directionality delimits group boundaries not only in visual gestalts, but also in musical gestalts. Melodically, for example, one hears a passage of individual notes as belonging together when the notes melodically rise or fall in a patterned way.

As stated above, the principle of good continuation asserts that one will perceive a group when the individual phenomenal objects seem to align together. This can be observed in the following musical gestalts:

5.2.1 Ascending and Descending scales

Consider the following musical gestalt, intended to be a musical equivalent of 5.1.2 above:



Not considering the principle of directionality, various patterns are possible, including:

5.2.2



5.2.3



5.2.4



Of course what once hears, due to the principle of directionality, is the following:

5.2.5



5.3 Directionality Principle in Prosodic Gestalts

In prosodic gestalts the principle of good continuation is also an important and frequent principle in the identification of group boundaries. Consider the following example, which demonstrates the principle of good continuation in prosodic gestalts:

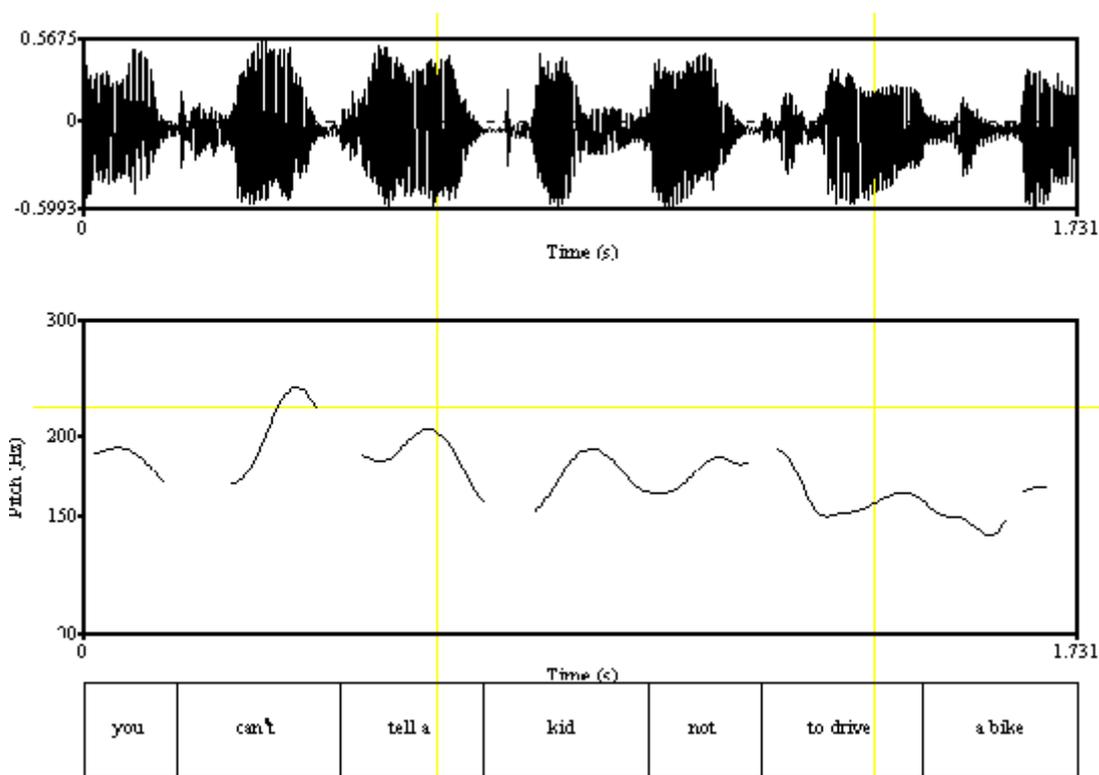
5.3.1 “You Can’t Tell a Kid Not to Drive a Bike”

Radio moderator Barbara Carlson tells a caller about the dangerous things that children do and explains that parents are helpless to protect their children against all accidents all of the time:

- 1 B: I can't tell you if it's the same one
- 2 but I can tell you that kids climb
- 3 and klids go out windows
- 4 and kids do all sorts of things .hh
- 5 and that when you have youngsters
- 6 um and you're living in an apartment house
- 7 and you do have railings
- 8 you got to watch them
- 9 you have to watch them
- 10 but what do you say to your child
- 11 you can't climb that tree Tucker
- 12 you can't do that Ann

13 I can remember when Tucker .hh
 14 <<*rising*> came home on his bicycle>
 15 <<*rising*> after cracking his tooth open>
 16 <<*rising*> by hitting a garbage can>
 17 → I mean you can't tell a kid not to drive a bike

The principle of directionality can clearly be heard in line 17, in which the pitch contour is manipulated by the speaker so that the pitches of the stressed syllables – which occur at isochronous intervals – fall in a consistent and patterned way. This can be seen in the following acoustic analysis:



Once a pattern is established, phenomenal events which continue the pattern will be perceived as part of the group. A phenomenal event which breaks that pattern will disrupt the directionality, thereby marking the end of the group. Consider, for example, the following list of numbers:

5.3.2 “KSTP Telephone Numbers”

Peter, co-moderator of the Barbara Carlson and Friends Show, gives the listeners the telephone numbers to call into the show.

→ P: six four six eight two five five
→ one eight-hundred
→ nine six two zero seven nine two
.hh star fifteen-hundred if you're a cellular one user
we have Mike on the line for you [Barbara
B: [Mike good morning

In line one Peter starts a list of numbers, which are the three telephone numbers for callers to call into the show. I will concentrate on the way Peter groups the numbers in lines 1 – 3. These lines consist of sixteen numbers: “six”, “four”, “six”, “eight”, “two”, “five”, “five”, “one”, “eight-hundred”, “nine”, “six”, “two”, “zero”, “seven”, “nine”, “two”. It will be seen that the principle of directionality operates to group these elements into three distinct groups of seven syllables, four syllables, and nine syllables respectively.

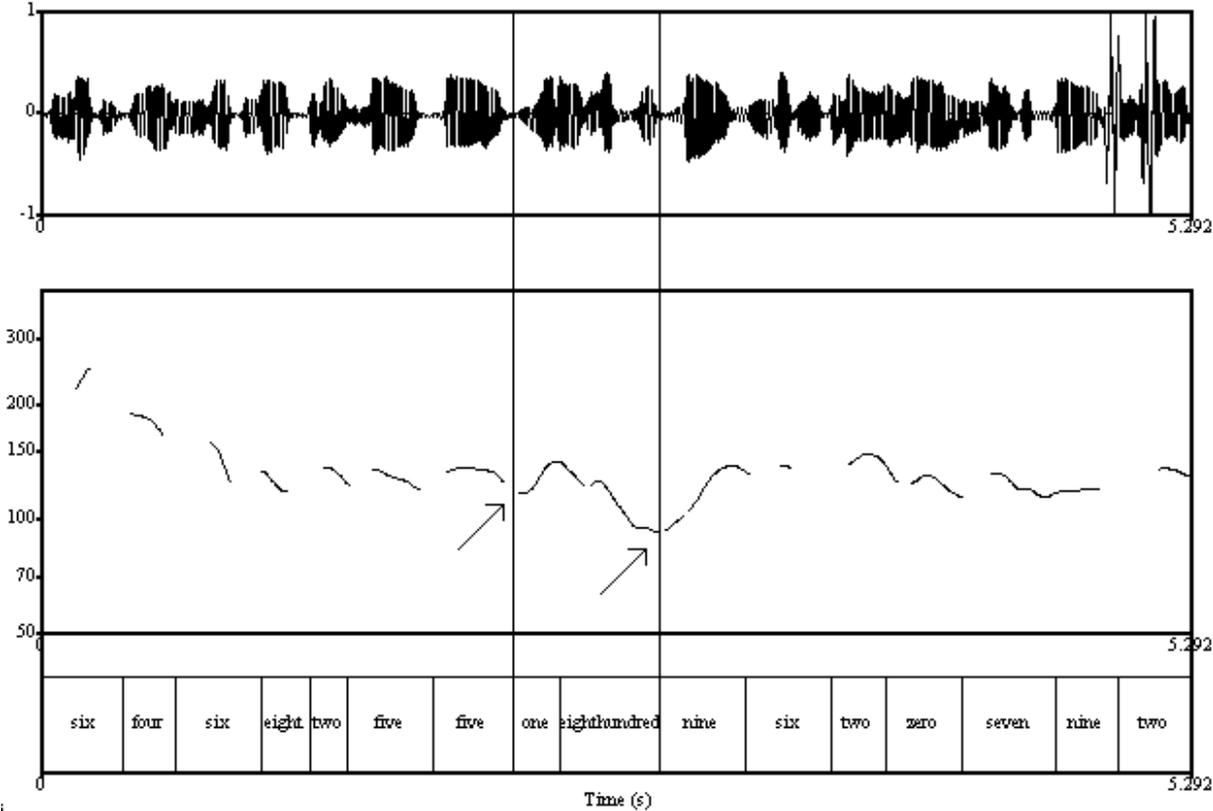
In line 1 Peter starts the list of numbers at approx. 250 Hertz. The next seven syllables steadily decline, ending at approx. 120 Hertz. No syllable starts at or lower than the preceding syllable and rises above the preceding syllable. They fall with smooth continuation.

This smooth continuation ends at the eighth syllable, “one”, which first starts below the pitch of the preceding syllable at approx. 100 Hertz. It then rises above the pitch height of the preceding syllable, ending at approx. 130 Hertz. The pitch movement of the eighth syllable above the pitch of the seventh syllable is heard as a disruption of the directionality of the preceding syllables. Consequently, one hears the first seven syllables as belonging to a group. A new group starts at the eighth syllable.

Just as the first group, the new group also distinguishes itself as its own group via the principle of directionality. The first syllable of this new group starts at approx. 150 Hertz. All of the following pitches steadily fall until the syllable “nine”, which, as the syllable “one” above, starts at the pitch of the preceding syllable but changes direction and rises above this last pitch. This marked change of direction causes the perception of the second group coming to a close and a new group beginning.

The pitch of the third group also rises to approx. 130 Hertz. The following pitches steadily decline. After the next nine syllables there follows audible in-breathing and a small pause, and another pitch step up. Due to the principle of directionality, these elements are similarly perceived as their own group.

The first syllable, by starting below and rising above the preceding syllable, breaks the principle of directionality, thereby giving the perception of a new group. In the following frequency analysis the two arrows point to those syllables whose pitch movement breaks the directionality principle:



There is, in fact, a practical reason for grouping the sixteen numbers as described above: the first group of numbers correspond to the local telephone number which callers may use; the second group of numbers correspond to the toll-free area code, and the third group of numbers correspond to the long-distance telephone number.

The following section will show how the principles of proximity, similarity and directionality can operate simultaneously in order to create intricate delimitations of the speech continuum. As in previous sections, I will illustrate the principles first with visual and musical examples, and then show how prosodic examples operate in a similar way.

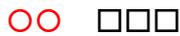
6 Compound Gestalts

The above examples showed how the principles of proximity, similarity and directionality operate within prosodic gestalts in order to create auditory groups. These principles, however, can also operate simultaneously in order to create “compound”, “nested”, and “complex” patterns. I call “compound gestalts” those phenomenal elements which are congruently grouped by two or more gestalt elements; I call “nested gestalts” those configurations in which there are 2 gestalts, A and B, such that B occurs within the larger gestalt A; I call “complex gestalts” those configurations in which two gestalts, A and B, form a pattern such that B starts after A starts but before A ends and ends after A ends. The following section 6 demonstrates compound gestalts using visual, musical and prosodic examples. In section 7 I will demonstrate nested gestalts, and in section 8 I will present complex gestalts, where the principles of similarity and proximity organize different stretches of speech in overlapping group structures.

6.1 Compound Visual Gestalts

To see how the principles of proximity and similarity can operate simultaneously in a compound gestalt, consider the following visual example:

6.1.1



In 6.1.1 the principle of proximity groups the first two shapes together and, following the spatial distance, the last three shapes together. The principle of similarity, congruent with the principle of proximity, groups the first two shapes together, which are circles, and the final three shapes together, which are squares. Congruent with the gestalt feature of shape is another feature, “color”. The first two shapes are not only circles, they are also red. The compound grouping structure of the above visual gestalt may therefore be represented as follows:

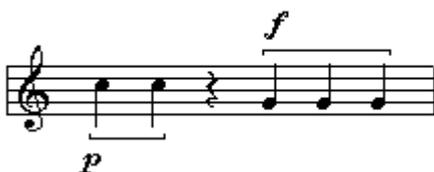
(1 (2 (3 XX 3) 2) 1) (4 (5 (6 XXX 6) 5) 4)

- 1 Principle of proximity
- 2 Principle of similarity (circle shape)
- 3 Principle of similarity (red color)
- 4 Principle of proximity
- 5 Principle of similarity (square shape)
- 6 Principle of similarity (black color)

As can be seen in the above example, the gestalt elements, in accordance with gestalt principles, congruently mark the same shapes. As such, they delimit what I term a compound gestalt.

6.2 Compound Musical Gestalts

An analogous compound musical gestalt would look as follows:



As in 5.1.1 gestalt elements congruently mark the same stretches of music according to both the principle of proximity and the principle of similarity. The first two quarter notes are perceived as a group and the final three quarter notes are perceived as a group because of the principle of proximity. They are also perceived as their own groups because of the low volume in the first group of notes, the high volume in the second group of notes, the same high pitch in the first group, and the same low pitch in the second group.

6.3 Compound Prosodic Gestalts

Just as there are compound visual and musical gestalts, there are also compound prosodic gestalts. Let us reconsider 4.3.1:

6.3.1 “We’re All Gonna Die”

Radio moderator Barbara Carlson, in a phone-in program concerning death and dying, talks to her radio audience:

```
1 B: we've all been born  
2   (.)  
3   <<1, pp> and we're all gonna die>  
4   (.)  
5   have you thought about it
```

As stated earlier, the principle of proximity separates the three lines. The principle of similarity forms a group of line 3: “and we’re all gonna die”, which is spoken at a lower pitch. Not only is the line spoken at a lower pitch register, it is also spoken very softly at pianissimo. The principle of similarity, which operates twice (volume and pitch register), and the principle of proximity congruently group the utterances into compound prosodic gestalts. The compound gestalts here form groups which coincide with the intonation unit.

7 Nested Gestalts

Consider the following visual example of a nested gestalt, which is set off by numbered brackets:

7.1 Visual Nested Gestalt

○ ○○ ○ ○○ ○ ○○ ○ ○○

The circles are grouped by the principle of proximity into an intricate nested pattern. The larger spatial distance in the middle (according to the proximity correlate) strongly creates the perception of two groups: six circles to the left and six circles to the right:

(₁○ ○○ ○ ○○₁) (₂○ ○○ ○ ○○₂)

Within the sub-group to the left, the proximity principle functions again to make two groups of three circles each:

(₁(₃○ ○○₃) (₄○ ○○₄)₁)

Within the subgroup of three circles, the proximity principle functions yet again to create one group which consists of one circle to the left, and one group of two circles to the right:

(₃(₅○₅) (₆○○₆)₃)

The nested grouping structure, in which one or more gestalts make up a larger gestalt, would look as follows:

$(_1(3(5X_5)(6XX_6)_3) (4(8X_8)(9XX_9)_4)_1) (2(7(9X_9)(10XX_{10})_7)(8(11X_{11})(12XX_{12})_8)_2)$

The circle which makes up group 5 is nested within group 3, which is nested within group 1.

Not only can the perception of nested gestalts be created through the principle of proximity, but also through the principle of similarity, as well as through a combination of both principles. Consider the following visual example, which displays the same nested gestalt structure as the above visual gestalt, but on the basis of both gestalt principles:

7.2 Visual Nested Gestalt

○□□ ○□□ ○□□ ○□□

The grouping structure, as in Example 6.1 above, can be seen as the following:

$(_1(3(5O_5)(6□□_6)_3)(4(7O_7)(8□□_8)_4)_1)(2(13(9O_9)(10□□_{10})_{13})(14(11O_{14})(12□□_{12})_{15})_2)$

Each individual group above is created by the following principles:

- 1) Proximity principle
- 2) Proximity principle
- 3) Proximity principle
- 4) Proximity principle
- 5) Similarity principle
- 6) Similarity principle
- 7) Similarity principle
- 8) Similarity principle
- 9) Similarity principle
- 10) Similarity principle

two measures, which have the same lower pitch range as the first two measures, are perceived as a group. What distinguishes these three groups is the principle of similarity: the first two measures are different from the second two measures, which are different from the final two measures. These three sub-gestalts may be grouped and bracketed as follows:

(1 da di du da di du da di du da di du 1) (2 da di du da di du da di du da di du 2) (3 da di du da di du da di du da di du 3)

Nested within these three gestalts are smaller gestalts, composed of three descending notes, which are grouped according to the principle of directionality. These smaller gestalts may be grouped and bracketed as follows:

(da di du)(da di du)(da di du)(da di du) (da di du)(da di du)(da di du)(da di du) (da di du)(da di du)(da di du)(da di du)

The intricate structure of this segment from “The Black Nag” with all of its nested gestalts grouped as a result of the principles of similarity and directionality, can therefore be depicted as in the following schemata, in which each note is represented as X:

[₁(₃XXX₃) (₄XXX₄) (₅XXX₅)] [(₆XXX₆) (₇XXX₇) (₈XXX₈)₁] [₂(₉XXX₉) (₁₀XXX₁₀) (₁₁XXX₁₁)] [₂(₁₂XXX₁₂) (₁₃XXX₁₃) (₁₄XXX₁₄)₂]

In the following section, I will demonstrate nested groups in prosodic gestalts.

7.4 Prosodic Nested Gestalts

For an example of prosodic nested gestalts let us reconsider 5.3.1, “You Can’t Tell a Kid Not to Drive a Bike:”

7.4.1 “You Can’t Tell a Kid Not to Drive a Bike”

Radio moderator Barbara Carlson tells a caller about the dangerous things that children do and explains that parents are helpless to protect their children against all accidents all of the time:

1 B: I can remember when Tucker .hh
2 <rising came home on his bicycle>
3 <rising after cracking his tooth open>
4 <rising by hitting a garbage can>
5 I mean you can’t tell a kid not to drive a bike

The principle of directionality causes the perception that lines 2, 3 and 4 each forms its own group. It therefore can be seen to have the following grouping structure: (₁came home on his bicycle₁) (₂after cracking his tooth open₂) (₃by hitting a garbage can₃). As each of these groups has the same melodic structure, each rising from approx. 220 Hz to 330 Hz, then all three sub-groups are part of a larger group that is formed through the principle of similarity. The three sub-groups, formed through the principle of directionality, are then nested within the larger group, formed through the principle of similarity. The grouping structure, then, can be seen as follows:

(₄(₁X₁)(₂X₂)(₃X₃)₄).

- 1) Principle of directionality
- 2) Principle of directionality
- 3) Principle of directionality
- 4) Principle of similarity

These are not the only nested gestalts which can be found, however. Grouped within the sub-groups 1, 2 and 3 are even smaller groups. Just as a small melodic phrase, for example, consists of individual notes which are nested within the phrase, a prosodic gestalt similarly consists of stressed and unstressed syllables which have certain pitches and which can be seen as groups in their own right according to the principles of proximity, similarity and directionality.

Another prosodic gestalt above the intonation unit which causes group perception is influenced by isochronous rhythm. The principle of directionality influences our perception such that when an audible rhythmic pattern develops in which stress syllables fall at equal intervals in time, the elements within that pattern are perceived as belonging together as a group. Isochronous rhythm is established when there are at least three equal beats: that is, when there are at least three stressed syllables which fall at equal distance in time to one another¹⁸. At the third beat, the principle of directionality takes over and the following beats are heard as belonging to the previous ones. Should the isochronous rhythmic pattern continue, all utterances whose stress syllables continue to fall within rhythmic pattern will be heard as belonging to the rhythmic gestalt. This perception will hold until the beat shifts and/or the isochrony breaks down.

In the above example (7.3.1), the three main accented syllables group the elements into the following global rhythmic gestalt according to the principle of directionality:

```

           came home on his
/1 bicycle after cracking his 1/
/2 tooth open by hitting a    2/
/3 garbage can                 3/

```

Within this global rhythmic gestalt, however, is another local rhythmic gestalt. Whereas pitch accent determined the global rhythm, syllable stress determines a more local rhythm, and forms a smaller rhythmic pattern which can be heard as nested within the larger global rhythmic pattern:

```

/1 came home on his 1/
/2 bicycle after    2/
/3 cracking his    3/
/4 tooth open by   4/
/5 hitting a       5/
/6 garbage can     6/

```

The phenomenon of one local rhythmic gestalt nested within a more global rhythmic gestalt I term “polyisochrony” (Flowe 1997). Conversation participants may interactionally exploit the nested local rhythm for different purposes than the global

¹⁸ Cf. Couper-Kuhlen 1990, 1993

gestalt. Consider the following example. V, A, B, G and O have just pulled up in front of B's apartment. B has been recording the conversation in the car and before he gets out, asks if there are any "last words" for the recording. The driver, V, begins:

7.4.2 "Möge Gott mit uns sein"

1 V: ehm
 2 möge Gott mit uns sein und ehm
 3 der himmlische Klang auf uns niederkommen
 4 A: <<all> auf Wiedersehen das Wetter>
 5 V: und uh
 6 möge
 7 A: <<f> ja doch>
 8 V: jeder eine gute Verdauung haben
 9 A: [yeach
 10 B: [<jingles keys>
 11 (2)
 12 A: warum muß echt immer in allem Verdauung vorkommen-
 Volker
 13 V: analhumor
 14 A: uhh

"May God Be With Us"

1 V: uh
 2 may God be with us and ehm
 3 the heavenly sound descend upon us
 4 A: <<all> goodbye the weather>
 5 V: and uh
 6 may
 7 A: <<f> alright already>
 8 V: each have good digestion
 9 A: [yeach
 10 B: [<jingles keys>
 11 (2)
 12 A: why does digestion have to come in everything Volker
 13 B: anal humor
 14 A: uhh

I argue that the playful modality is maintained in the global isochronous gestalt, established collaboratively by V and A in lines 1 – 9. In the local isochronous gestalt, however, A is trying to usurp V's established rhythm by establishing a quicker isochronous rhythm. She is trying to throw him off rhythm in order to make him stop. This interpretation is further supported in lines 4 and 7. In line 4 A says "auf Wiedersehen das Wetter", which is what newscasters in Germany say at the end of

the news program. The playful way of trying to force V to stop fits in the playful modality. In line 7, she is more direct and says “ja doch” (“*alright already*”).

The global rhythmic gestalt can be depicted as follows:

```

1      /ehm                               /
2      /möge Gott mit uns                 /
3      /sein und ehm                       /
4      /(.) der                            /
5      /himmlische Klang auf uns          /
6      /niederkommen auf wiedersehen das /
7      /Wetter                             /
8      /und uh                             /
9      /(.)                                 /
10     /möge ja doch jeder eine           /
11     /gute Verdauung haben              /
12     /[yeach                             /
13     [jingles keys
```

Here V speaks slowly, and the beats of the isochronous gestalt occur at approx. 1 second intervals. In line 6 A takes V's stressed beat within the global isochronous gestalt and uses it as the first beat of a local isochronous gestalt in which the beats occur twice as fast as in the global gestalt. This can be seen in the following:

```

/niederkommen auf /
/Wiedersehen das /
/Wetter           /
```

Rather than take over the new quicker rhythm, however, V continues his established global rhythm in line 8 in the rhythm transcription above. This is followed by a hesitation particle and a silent beat. In line 10 he continues the global isochronous gestalt with “möge”. A again tries to quicken the rhythm, and, as in line 6 above, she uses V's stressed syllable on “möge” as the first beat of a new local rhythmic gestalt, which is twice as fast: “ja doch”. This time, V adopts the local gestalt and finishes his utterance: “jeder eine gute Verdauung haben”. The local gestalt can be seen as follows:

```

/möge           /
/ja doch        /
/jeder eine     /
/gute          Ver- /
/dauung haben   /
```

/[yeach
[jingles keys

This quicker rhythm still fits in the global isochronous gestalt:

/und uh /
/(!) /
/möge ja doch jeder eine /
/gute Verdauung haben /

A's "yeach" and the key jingling are shown as part of the global pattern on p. 64.

This excerpt illustrates that a nested gestalt can have a different (even contradictory) function compared to the more global rhythmic gestalt in which it occurs. In this example, the playful modality is maintained by the global rhythmic gestalt, and indeed, one can continue to tap the beats even when the local rhythmic gestalt is used to throw V off, which would potentially end the sequence. The speakers are playing with prosody: A "attacks" V with a quick local rhythm, V adapts the rhythm, but is still able to maintain the global isochronous gestalt.

The above examples demonstrate the complicated patterns that may arise out of compound and nested gestalts. There is still, however, another kind of structure, which I will call "complex gestalt". The following section will explain "complex gestalts" using visual and musical examples. Prosodic examples of complex gestalts will then be offered.

8 Complex Gestalts

The above sections showed how the principles of proximity and similarity can operate in congruent ways to form compound groups and how groups can occur within other groups to form nested gestalts. The principles of proximity and similarity can also operate in a more intricate way so that the stretches of speech are delimited differently, thereby forming complex gestalts. As stated above, two gestalts A and B form a complex gestalt if B starts after A does, and if B also ends after A does. Such grouping structure would therefore be as follows:

$(_1X(_2X_1)X_2)$

As in the other sections, I would like to demonstrate complex gestalts using visual, musical, and finally, prosodic gestalts.

8.1 Complex Visual Gestalt

Consider the following visual gestalt:

8.1.1



According to the principle of proximity, the first three shapes are grouped together, and the final two shapes are grouped together. According to the principle of similarity, however, the first two shapes are grouped together (as they have the identical circle shape) and the final three shapes are grouped together (as they have the identical square shape). Furthermore, the principle of similarity operates again, grouping the first four shapes together (which are in red), and the final shape is perceived as its own group (which is in blue). The complex grouping of the above visual gestalt is therefore as follows:

$(_1(3(5XX_3)(4X_1)(2X_5)(6X_2)4)6)$

- ₁ Principle of proximity,
- ₂ Principle of proximity,
- ₃ Principle of similarity (circle shape),
- ₄ Principle of similarity (square shape),
- ₅ Principle of similarity (red color),
- ₆ Principle of similarity (blue color).

Such complex grouping structures can also be observed in temporal musical gestalts, as demonstrated below.

8.2 Complex Musical Gestalt

Consider the following musical phrase:

8.2.1



This musical temporal gestalt has the same complex grouping structure, and the principles of proximity and similarity operate in the same way to give this example the same grouping structure as the visual gestalt. The first three quarter notes are perceived as a group because of the principle of proximity, as are the final two notes. The first two notes are perceived as a group because of higher pitch; the last three notes because of lower pitch. Low volume delimits the first four notes; high volume delimits the final note. As seen in the above two examples, different phenomenal elements can be grouped differently depending on the gestalt element involved. Like the visual gestalt 8.1.1 above, 8.2.1 has the same complex grouping structure:

$$({}_1({}_3({}_5XX_3))({}_4X_1))({}_2X_5)({}_6X_2)_4)_6$$

- ₁ Principle of proximity,
- ₂ Principle of proximity,
- ₃ Principle of similarity (high pitch),
- ₄ Principle of similarity (low pitch),
- ₅ Principle of similarity (low volume),
- ₆ Principle of similarity (high volume).

8.3 Complex Prosodic Gestalt

Complex grouping structure can also be seen in prosodic gestalts, where the boundaries between different prosodic structures are established in incongruent ways.

As an example of complex grouping in a prosodic gestalt, consider the following:

8.3.1 "Buick Safety"

Radio moderator Paul Harvey, in a radio program that weaves news items and advertisements together, presents a Buick car advertisement.

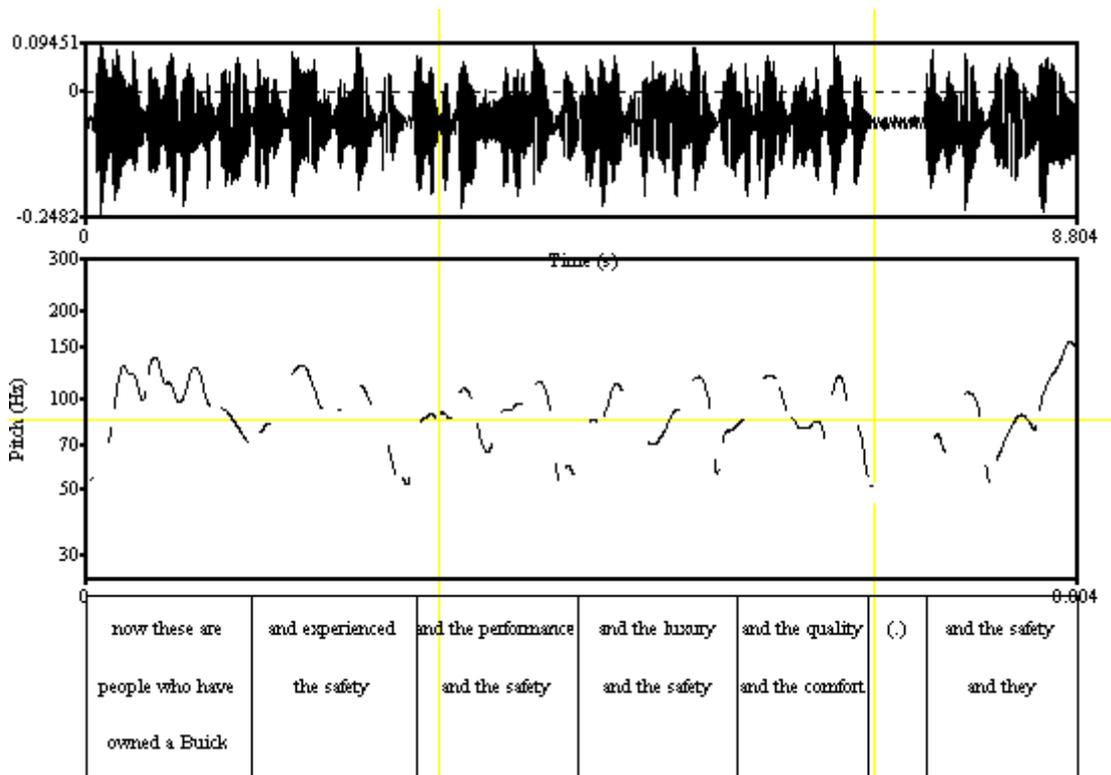
```
1   now these are people who have owned a Buick and-
   experienced the
2   /safety and the per-/
3   /formance and the  /
4   /safety and the    /
5   /luxury and the    /
6   /safety and the    /
7   /quality and the   /
8   /comfort           /
9   /(.) and the       /
10  /safety and they   /
11  /.hhh
12  are most likely of all car owners to make certain their-
   next car is also a Buick
```

As stated above, the principle of proximity separates the speech at line 9, where a silent beat occurs. But the gestalt is even more complex. Starting in line two, and continuing to line 10 there is a particular melodic pattern, where Harvey manipulates the pitch of the elements so that the pitch of the first syllable is approx. 120 Hertz and the pitch of the second syllable is at approx. 70 or at approx. 55 Hertz. This melodic pattern continues after the silent beat, where, in line 10, "safety" is repeated with the first syllable at approx. 118 Hz and the second syllable at approx. 70 Hertz.

The gestalt is yet even more complex. The isochronous rhythmic pattern also marks the speech starting in line 2 and continuing to line 11, where the audible breathing is heard to fall directly on the beat set up by the speaker. Following the breathing, the isochronous rhythm breaks down.

The following acoustic analysis clearly shows the pause in the wave form, as well as the speaker's orientation towards pitch height and valleys in the frequency analysis.

8.3.1.1 Acoustic analysis of "Buick Safety"



The syllables "and the safety" in lines 9 and 10 are therefore considered complex because according to the proximity principle, they belong to the phenomenal items following them; melodically (as well as semantically), they belong to the phenomenal items preceding them.

The complex gestalt of the sequence therefore looks as follows:

(₁now these are people who have owned a Buick and experienced (₃(₄the safety and the performance and the safety and the luxury and the quality and the comfort₁) (₂and the safety₃)₄) and they₂)

- ₁ Proximity principle
- ₂ Proximity principle
- ₃ Principle of similarity (isochronous gestalt)
- ₄ Principle of similarity (melodic gestalt)

The result of the different grouping patterns is a feeling of surprise at line 10. After the strong perception of closure, indicated by the long pause and the steep falling pitch on the last syllable, the pattern unexpectedly continues melodically and rhythmically. The long pause in line 9 is then heard as a silent beat within the more global isochronous gestalt.

9 Summary

The above has claimed that prosodic structures are perceptible gestalts which, from a gestalt theoretical perspective, are similar to visual and musical gestalts. There are, of course, other linguistic and non-linguistic levels which are also an important part of the interaction: syntax, semantico-pragmatics, visual parameters, etc. Considering turn-transition places, Auer asserts that an utterance may be redundantly marked on a number of different levels. According to Auer: “syntactic, prosodic, semanticopr pragmatic and visual parameters “may (and indeed often do) co-occur (i.e. there is often a certain amount of redundant signaling); and their interactive effect cannot be taken back or ‘interactionally denied”” (1996:58). I have concentrated on the delimitation of prosodic gestalts, but the object of study becomes even more complex, however, when one additionally considers other levels.

Although not true gestalts¹⁹, syntactic and semantic structures can also group the speech continuum. One can imagine, for example, complex patterns in which syntax groups the speech continuum in one way, prosody in another way. Reconsider the following example (4.3.1 above), which was used above as an example of a compound prosodic gestalt:

- 1 B: we've all been born
- 2 <<l, pp> and we're all gonna die>
- 3 have you thought about it

¹⁹ Although syntactic and semantic structures display some similarities to prosodic Gestalts, I reserve the term “gestalt” to refer only to perceptible structures which are governed by gestalt principles, such as prosodic gestalts.

Above, I argue that this sequence can prosodically be grouped as follows:

(x)((x))(x)

Syntax and lexical repetition, however, create a different grouping, as demonstrated in the following:

[[x][x]][x]

Considering these linguistic levels simultaneously results in the following structure:

[[(x)] [((x))]] [(x)]

A syntactic structure, therefore, does not have to be grouped equivalently with a prosodic gestalt. List intonation, for example, might be used to group certain parts of an utterance, even though semantically and even syntactically, the items are not a list. Were it not for prosodic grouping these items would not be perceived as belonging together at all. Such an intricate structure over linguistic (and non-linguistic) levels – though not the focus here – is still deemed important and could be the focus of future research.

Chapter 2: The Interactional Relevance of Compound, Nested and Complex Gestalts at, below, and above the Intonation Unit Level

The following chapter has a dual purpose: It demonstrates the interactional relevance of the kinds of grouping structures proposed in chapter 1 by showing interactants orienting towards compound, nested and complex gestalts within discourse. Additionally, it shows that there are grouping structures at, below and above the traditional intonation unit. The chapter is divided into three sections: in the first I will show the interactional relevance of grouping structures at the level of intonation phrase, in the second I will show the interactional relevance of grouping structures below the intonation phrase, and in the third section, I will show the interactional relevance of grouping structures above the intonation phrase.

1 Introduction

The previous chapter proposes the existence of intricate prosodic groupings at, below and above the level of intonation unit. I claim that the possibilities interactants have for more intricately grouping the speech continuum are far greater than current prosodic research provides for. The intonation unit, an important linguistic construct available for speakers, is seen as one kind of grouping among others.

The following chapter intends to show that such intricate grouping is not only a scientific construct but is rather interactionally relevant to the conversation participants themselves. Speakers who are showing an orientation to the proposed phenomena are simultaneously giving evidence that these phenomena are interactionally relevant. To show the relevance of the proposed gestalts I will present examples in which speakers orient towards the complex, nested or compound gestalt by presenting 1) the repetition of a gestalt structure, 2) a gestalt which is picked up and repeated by another speaker, 3) a collaborative completion by two or more speakers of a compound, complex, or nested gestalt. The repetition of a gestalt (either by a single speaker or by multiple conversation participants) and collaborative

completion of gestalt structures indicate that the patterns are, in fact, not random but are relevant within discourse.

According to Hutchby and Wooffitt, one of the “core questions” for a CA analyst when approaching data is “how do participants demonstrate their active orientation to this (interactional) business” (1998:104). The orientation speakers show for an interactionally relevant phenomenon can be expressed in a number of ways. Most simply, they recognize the phenomenon and verbally express it in a meta-communicative comment. While this is often the easiest evidence to identify and present, it is not the most frequent.

More common is for speakers to display an orientation towards the phenomenon in more subtle ways. Ono/Thompson (1995) list two ways in which speakers show their orientation towards a grammatical construction or schema:

We simply wish to show that the conversational data are rich with evidence in favor of the type of schemas that Langacker has argued for. We will discuss two types of evidence: first (...) speakers seem to attend to schemas even across utterances, either their own or another speaker's, which are not part of those schemas. Second (...) utterances jointly constructed by two conversational participants provide strong evidence for constructional schemas.

(Ono/Thompson 1995:222)

The second kind of evidence offered by Ono/Thompson, co-construction, is divided into two kinds:

Our data reveal two types of co-constructions. The first is the situation that obtains when one speaker adds to a unit spoken by another speaker, and this addition is accepted by the first speaker. In such cases, although there may be no evidence that the first speaker had ‘planned’ to say what the second speaker adds, there is evidence that they are both content with the jointly produced result. (...) The second type of co-construction we found is exemplified by data in which there is evidence that the second speaker, rather than adding something to another person's utterance, finishes what the first speaker had started to say.

(Ono/Thompson 1995:227-228)

Although Ono/Thompson present evidence for syntactic schemas, the kinds of interactional evidence which they cite can equally be found for the kinds of prosodic gestalts I propose. A prosodic pattern that is attended to across utterances is seen as evidence for the relevance of the gestalt, as well as a prosodic pattern that is formed by a first speaker and then picked up and repeated by one or more speakers.

The phenomena which will be shown to be interactionally relevant here are compound, nested and complex gestalts at, below, and above the level of the intonation unit. In the first section I will present compound gestalts, i.e. gestalts that coincide with the intonation unit. In the second section I will present nested gestalts below the level of the intonation unit, and in section three I will present complex gestalts above the intonation unit.

2 Prosodic Gestalts at the Level of the Intonation Phrase

Most frequently, speakers divide the speech continuum into groups at the level of the intonation phrase. Occasionally, certain intonation contours at the intonation phase level become stylized, and an orientation towards them is displayed by the interactants themselves.

Consider, for example, the following excerpt, taken from a telephone conversation between the two intimates, Jane (J) and Ben (B). Earlier in the conversation, Jane tells Ben that her period is late, but a pregnancy test has shown the results to be negative. The following excerpt occurs about 20 minutes later in the conversation. Jane (J) and Ben (B) discuss Ben's possible reaction if the results had been positive:

2.1 “No Responsibility”

- 1 B: aw honey
- 2 that would just like ruin your whole idea of who I am
- 3 hunh
- 4 J: aw
- 5 B: it would just like
- 6 I'd just blow up everything that we had

7 if I
 8 <sniff>
 9 if I was that cruel
 10 J: if you just refuted it
 11 B: if I said ↑nO: ↓wA:y
 12 J: ↑nU:n ↓hU:nh
 13 I'm not taking ↑Any responsi↓bIlity
 14 B: if it's ↑mIne ↓thEn
 15 it's ↑yOUr fault for getting ↓prEgnant
 16 J: <<piano <oh God
 17 B: wo ha ha ha ha no

On the semantic and syntactic levels, Ben gives the consequence of his behavior in line 6: “I’d just blow up everything that we had”. He then offers the conditions for this consequence starting in lines 7 and 9: “if I <sniff> if I was that cruel”. Jane picks this up in line 11: “if you just refuted it.” In line 11 Ben states another condition, “if I said ↑no ↓way”, which shows a shift into direct speech. Following this Jane stops using the conditional, but brings the conditional world to life via a deictic shift. She speaks in Jeff’s voice in the present tense in the conditional world they have just been imagining: “nun hunh / I’m not taking any responsibility.” Jeff takes over, speaking in his voice in the imaginary world in lines 14 and 15: “if it’s mine then / it’s your fault for getting pregnant”. This is followed by Jane’s shift back into the “real world” with her exclamation: “oh God”.

Starting in line 11 and continuing through line 15, a particular intonation contour is presented and repeated five times, three times by Ben, twice by Jane. For this contour, the stressed syllable of the head and the nucleus are prosodically formatted such that the first stressed syllable of the head rises to a certain pitch. This pitch is sustained throughout the head before falling sharply at the nucleus to a harmonic interval heard as a musical fifth below the pitch of the head.

B: if I said ↑nO: ↓wA:y
 J: ↑nU:n ↓hU:nh
 I'm not taking ↑Any responsi↓bIlity
 B: if it's ↑mIne ↓thEn
 it's ↑yOUr fault for getting ↓prEgnant

The repetition by both interlocutors shows a strong orientation to the gestalt pattern, created by the principle of similarity. Indeed the gestalt coincides with the imaginary world presented by Ben and Jane. Ben's first move into direct speech in line 11 coincides with the intonation contour as described above. This contour continues until line 16, when Jane shifts her prosody with her utterance "oh God", suddenly speaking piano and without the stylized contour of the five previous intonation phrases. The prosodic shift coincides with a shift out of the imaginary world and back to the real world. Ben's response in line 17, in which he says "oh no" while simultaneously laughing, indicates that he, too, is back into the real world, responding to Jane's move out of the imaginary world. Had Jane continued to orient to the gestalt, reproducing it again in line 16, she would arguably have been reproducing Jeff's voice in the imaginary world rather than giving a display of her own interjection in the "real world"²⁰.

For another example of orientation towards a compound group, which is the size of the intonation phrase, consider the following example:

2.2 "The Preacherly Vine"²¹

In this excerpt, Elder Otis Jones (E) preaches to a congregation (C).

- 1 E: well hAllelUjah;
 2 prAise the LO:rd;
 3 C: <response>
 4 E: we thAnking the LOrd for bEing here this A:fternOOon;
 5 C: <response>
 6 E: that we all can prAise GOD together;
 7 C: <response>
 8 E: I belIEve you know in the fIfteenth chApter of SAInt=
 =JOhn?
 9 C: <response>
 10 E: ahh and Jesus said I am the trUe vIne-
 11 C: <response>
 12 E: and my Father is the husbandmAn;
 13 C: <response>
 14 E: Every brAnch in me that (...) bringeth frUIt?

²⁰ Cf. Klewitz/Couper-Kuhlen (1999).

²¹ Recording from "Religious Music: Congregational & Ceremonial". Recorded Sound Section, Music Division, Library of Congress, Washington, D.C. Library of Congress Catalog Card Number 75-750854.

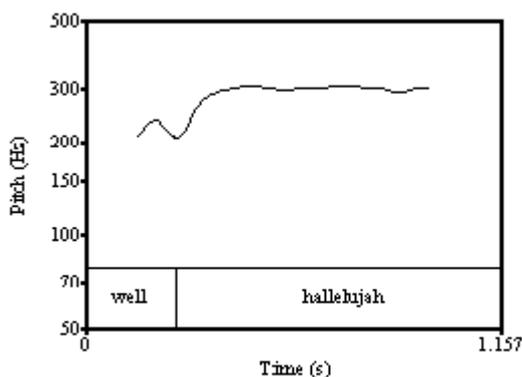
15 C: <response>
 16 E: mY hEAvenly FAtHer pURges It;
 17 C: <response>
 18 E: that it mIght brIng fOrth mOre frUIt.
 19 C: <response>

In this excerpt, taken from an African-American Christian worship service, the preacher (E) repeats a global gestalt (principle of similarity) which is created equally by a very stylized intonation contour [see Chapter 3]. Furthermore, this gestalt is oriented towards by the congregation (C), which gives certain recipient signals that provide evidence for the relevance of the gestalt.

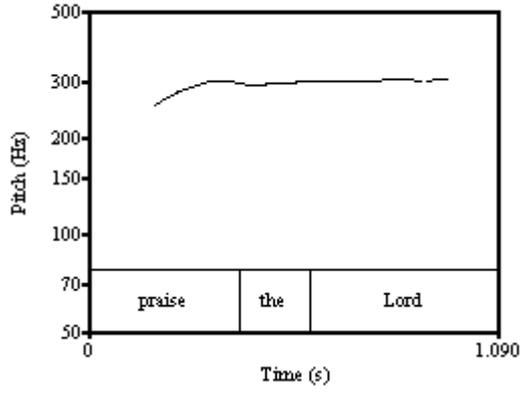
The local gestalt which gets established is heard to have a very stylized pitch pattern. Starting with the first stressed syllable of the intonation phrase, the speaker glides to a pitch of approx. 300 Hertz, and stays at that pitch until the nucleus, which may rise dramatically (lines 8 and 14), fall (line 18), or, as is most frequently the case, retain the pitch (lines 4, 6, 10, 12, 16). Whether or not the pitch rises, falls, or stays the same, the last syllable of the intonation phrase is always lengthened.

An acoustic analysis of the fundamental frequency of the first four lines captures our impression of a stylized melodic gestalt:

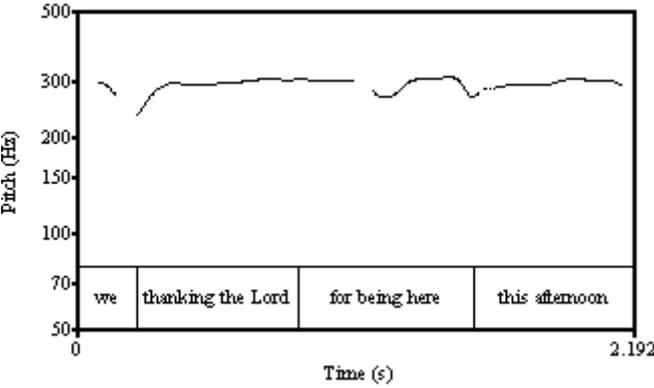
Line 1:



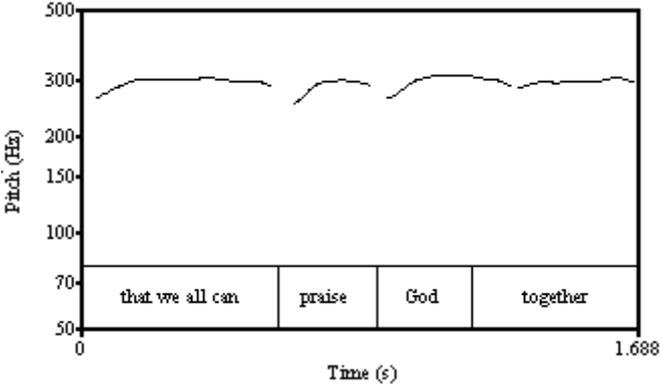
Line 2:



Line 3:



Line 4:



The above frequency analyses show that the speaker highly stylizes his utterances²². The pitch of each contour – not only in these lines but also throughout

²² See chapter 3: "A Gestalt Approach to the Highlighting and Stylization of Prosodic Structures".

the sequence – consistently stays around 300 Hertz. This can easily be seen in the above four examples.

Starting in line 3, and heard after every intonation phrase by the preacher, the congregation gives a collaborative response. Some people are heard to say “amen”, “hallelujah”, etc. Evidence that the preacher’s intonation phrase is interactionally relevant is the precise placement in time of the congregational response.

Further evidence that the congregation is orienting to the intonation phrase can be heard in line 14. In this line, the preacher stops in the middle of the established intonation phrase and pauses for half a second:

E: Every brAnch in me that (...) bringeth frUIit?
C: <response>

Not a single member of the audience responds during this pause, which gives strong indication that they are orienting towards a particular prosodic gestalt, and that the gestalt is not yet complete. They are waiting for the nucleus, which occurs after the pause and with strong upwards pitch movement. This is then followed by the audience’s response.

3 Nested Prosodic Gestalts below the Level of the Intonation Phrase

Whereas the above examples demonstrate participants’ orientation towards compound gestalts that are congruent with the intonation phrase, the following will demonstrate participants’ orientation to nested structures which are smaller than the intonation phrase.

Consider, for example, the following excerpt. In an informal conversation between two cousins, Alina and Lenore, Alina tells Lenore about meeting her brother-in-law, Antonio:

3.1 “Ugliest Set of Shoes”

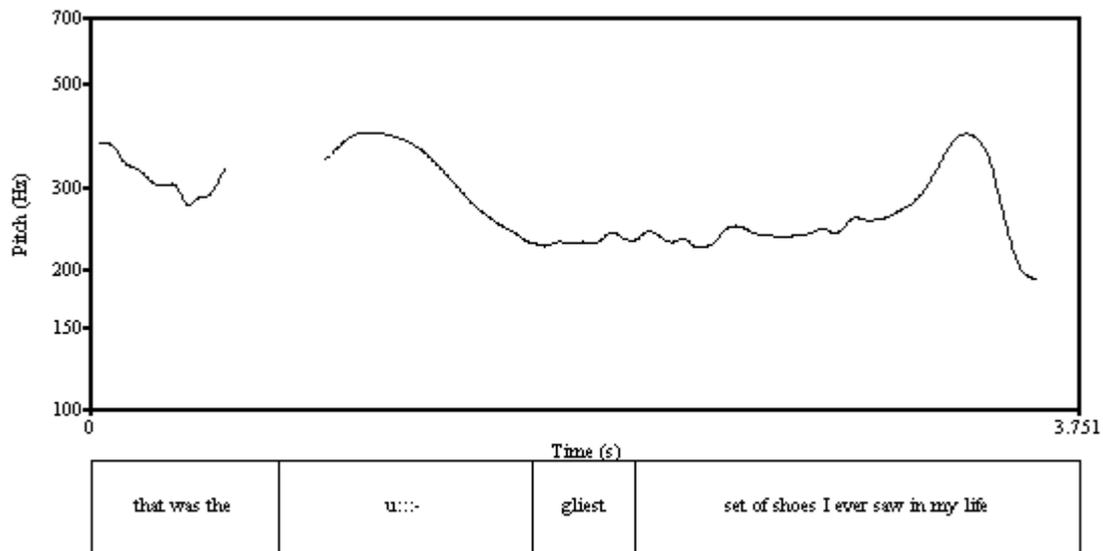
1 A: Antonio had on white shoes
(.)
2 L: [I'm sorry
3 A: [I don't care if he is Italian
4 L: [ha
5 A: [I don't care
6 that was the (.) <u::gliest set of shoes I ever saw
=in my life>
7 ha ha ha ha
8 L: .hhh

In line 1, Alina tells her cousin that she has met Antonio; moreover, she describes what he was wearing: white shoes. Following a pause, Lenore, in overlap with Alina, says “I’m sorry”. It may be argued that her contribution here is a kind of negative assessment for something which has gone wrong, in this case Antonio’s bad taste in shoes. In any case, Lenore continues to set up a truly negative assessment in line 3 when she says, “I don’t care if he is Italian”. This is followed by a single laugh particle by Lenore before Alina continues with “I don’t care” in line 5. Alina is building up a contrast, which finally comes in line 6: “that was the (.) ugliest set of shoes I ever saw in my life”. This is followed by extremely loud, rhythmic and melodic laughter by Alina and loud audible breathing by Lenore.

Within the intonation phrase of line 6: “that was the (.) ugliest set of shoes I ever saw in my life” is a nested stylized gestalt. Starting with the first stressed syllable “u::gliest” and ending with the nuclear syllable “life”, one hears a stylized prosodic gestalt nested within the intonation phrase. Following a micro-pause after the anacrustic syllables “that was the” Alina reaches with the first syllable of “u::gliest” a pitch of 400 Hertz. This exact pitch is reached again with the nuclear syllable, “life”. All the pitches between “u::gliest” and “life” are kept at a constant level of approx. 235 Hertz.

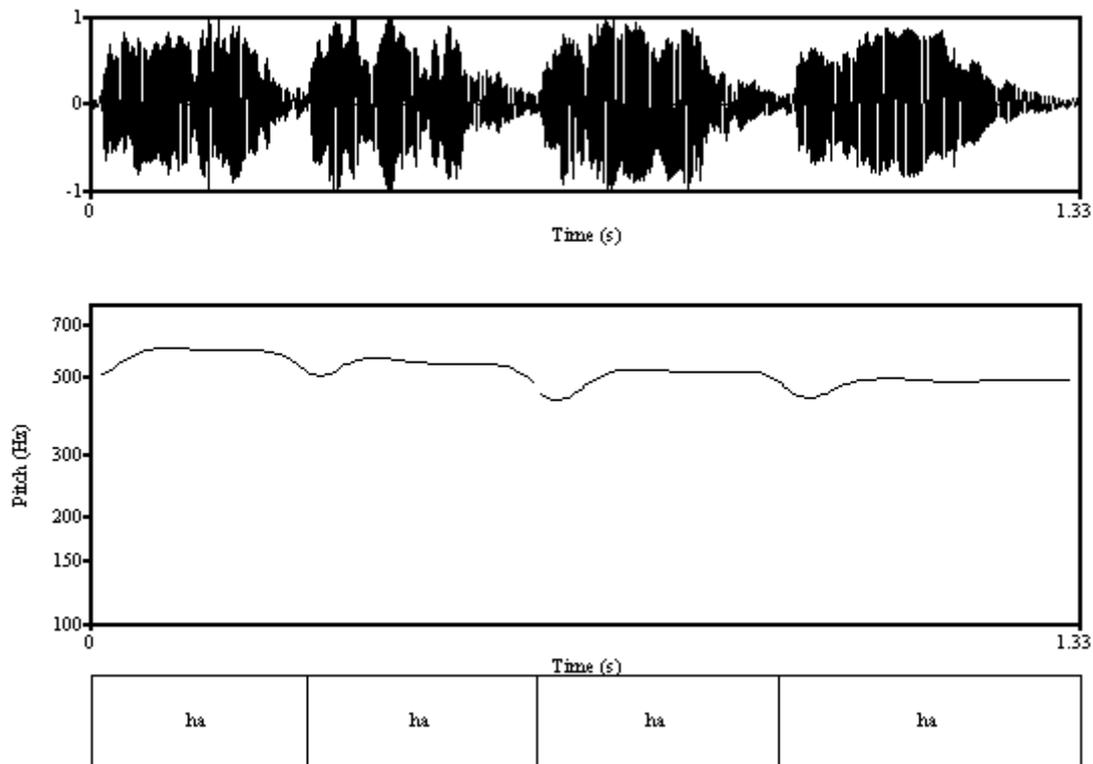
An acoustic analysis of fundamental frequency supports the auditory impression:

3.2



Evidence of orientation to the stylized nested prosodic gestalt is Alina's own highly stylized laughter in line 7. Each laugh particle is rhythmically and melodically formed. The pitch of each laugh particle falls in a patterned way, as shown by the following acoustic analysis:

3.3



It could be argued that Alina's highly stylized laughter would be out of place if the previous utterance were not likewise highly stylized. Though she does not use the same prosodic design that she uses in line 6, she does follow prosodic stylization with prosodic stylization. Moreover, her laughter, with a specific rhythmic and pitch contour, as shown by the above acoustic analysis, is a definite auditory gestalt. As such, the laughter represents a kind of phenomenon – a non-verbal auditory event – which has been neglected by traditional prosodic research.²³

Alina begins the stylized gestalt in line 6 not at the beginning of the intonation phrase, but rather in the middle of it. The stylized acoustic gestalt begins with the lexical word “ugliest”, which underlines the paradox which can be summed up as follows:

Italians have good taste in shoes

Antonio is Italian

Antonio does NOT have good taste in shoes.

By starting her stylized prosodic gestalt at the lexical word which marks the semantic contrast and not with the anacrustic syllables “that was the”, Alina is using prosody to draw attention to the point of greatest semantic contrast. As such, it makes the assessment, “ugliest shoes” even stronger than it would have been had there been no prosodic stylization.²⁴

For another example of participants' orientation to a nested gestalt below the intonation phrase, consider the following example:

3.4 “That Made Me Mad”

```
1 M: I couldn't find it lAst time I was in there
2   that that made me [ma:d
3   ?:                               [<knock>
4 K: [<<low> ma::d>
```

²³ Although laughter has been neglected in much prosodic research, it has been the focus of research from an interactional perspective. Cf. Jefferson (1979, 1984), Jefferson, Sacks, Schegloff (1987), Phillip J. Glenn (1989, 1991), Emanuel Schegloff (1993), Laura Gavioli (1995), *et al.*

²⁴ The ability of prosodic stylization to upgrade or downgrade assessments is taken up in Chapter 5, “Prosodic Stylization and Assessment Sequences”.

5 [*knock*
 6 W: I was sO:: mA::d
 7 M: I was [ma:d in the [hEA:d
 8 W: [dO: do [dO: do
 9 ?: [<knock> [<knock>
 10 W: don't you remEmber that
 11 K: [unhunh
 12 ?: [<clap clap>
 13 K: how mad were you
 14 M: oh fairly

In line 2 Marci states, “that made me ma::d”. Compared to the other syllables in the intonation phrase, “mad” is low in pitch and lengthened. Occurring simultaneously with the lengthened syllable is a knock.

In line 4 Kendra shows her orientation to this smaller structure within the intonation phrase. She repeats the syllable: “ma::d”, which is realized not only with lengthening and a lower pitch, the syllable is even more lengthened than Marci’s and deeper. As with Marci’s “mad”, there occurs a knock simultaneously with the syllable. Kendra not only repeats Marci’s stylized gestalt, she adds to it, making it even more stylized.

Unlike a speaker’s vocal contributions, which an analyst can match to a particular speaker with reasonable accuracy due to individual voice quality, the producers of non-vocal sounds, such as taps, claps, knocks, etc. are more difficult to identify. For the interactants, however, such non-vocal sounds are still important aspects of the auditory gestalt. These non-vocal sounds, important though they are for the interactants as part of their utterance, have only rarely been considered in past prosodic research²⁵.

In Marci’s contribution a knock is heard at the stylized nucleus. In Kendra’s repetition of Marci’s “mA:d” a second knock is also heard. This is interpreted as not mere coincidence, and in the following unfolding discourse, further non-vocal phenomena play a specific role in the creation of even more stylized prosodic gestalts. Kendra’s contribution specifically shows her orientation, not to the entire intonation phrase, but to only part of it.

²⁵ One notable exception is Erickson (1981, 1992).

Not only does Kendra repeat “mA:d” in line 4, Wendy also repeats it. Just as Kendra’s “mA::d” is heard as more stylized than Marci’s due to greater lengthening and deeper pitch, Wendy forms her contribution so that it is even more stylized than Kendra’s “mA:d”. In line 6 Wendy says “so::: mA:::d”, which is spoken with low pitch and even more lengthening than the previous utterances.

In the following lines, speakers develop the stylized gestalt even more, not only lengthening their syllables, but actually singing them. In line 7 Marci sings “I was mA:d in the hEA:d”, which perfectly overlaps with Wendy’s “dO: do dO: do” in line 7. The anacrustic syllables “I was” are realized on a single tone and the stressed syllable “mad” is heard as a harmonic fifth above it. Rhythmically, “I was” is realized twice as fast as “mad”. This pitch and rhythmic pattern is repeated with “in the head”. Written out in music notation, Marci’s contribution can be transcribed as follows:

3.5



Not only is Marci’s contribution highly stylized, it occurs in overlap with Wendy’s contribution. Rhythmically, Wendy’s “dO: do dO: do” occurs simultaneously with Marci’s stressed syllables “mad” and “head”. Furthermore, a knock is also heard precisely at these syllables. In music notation, the intricate pattern is heard as follows:

3.6

Marci:
(line 7)
I was mad in the head

Wendy:
(line 8)
do do do do

?:
(line 9)
* *

In this example, the speakers are heard to orient towards a nested gestalt, and then to develop it into a highly complex stylized gestalt. The example shows orientation to a gestalt both below the intonation phrase (Marci's original "mad" in line 2) and to the highly intricate auditory gestalt in lines 7, 8, and 9, which demonstrates a gestalt structure above the intonation phrase.

For another example of a nested gestalt, consider the following sequence:

3.7 "In a Voice Like This"

In an informal conversation, Alina (A) tells her cousin Lenore (L) about visiting another family member, Jan, with C's mother.

- 1 A: it's amazing
- 2 no Jan talked the whole time <<high, fast, loud> in a=
=voice like this
- 3 h[i::: Mar Ali:::na I'm so happy to see you>
- 4 L: [ha ha ha
- 5 A: I'm going
- 6 ha ha ha
- 7 Go::d
- 8 turn the volume down
- 9 let me outta here>
- 10 L: ha ha ha

In this excerpt, Alina begins telling her cousin about her unfortunate visit with another family member, Jan. In line 2 Alina says: "Jan talked the whole time in a voice like this". This syntactically complete line is simultaneously one intonation

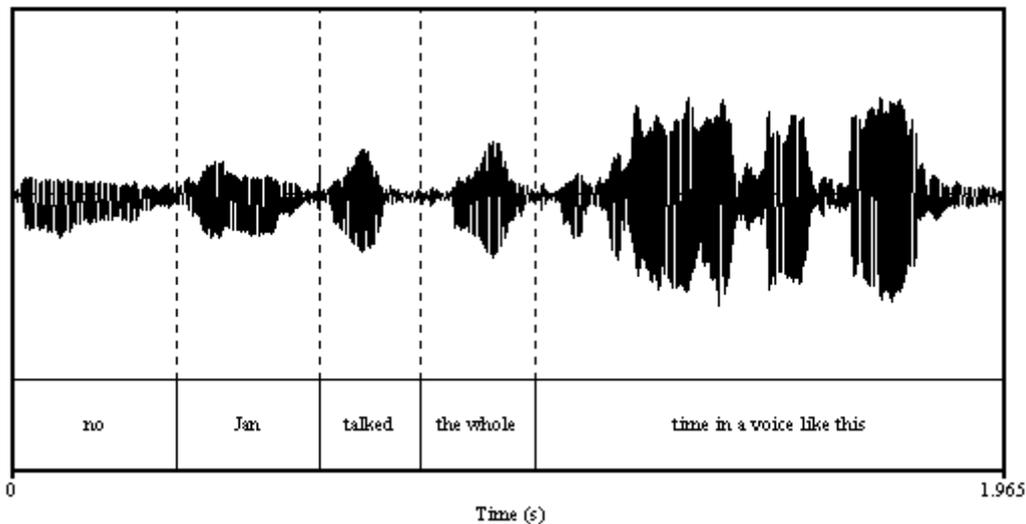
phrase, but two prosodic gestalts. The structure of the intonation phrase can be seen in the following diagram:

Prehead	Head	Nucleus
Jan	talked the whole time in a voice like	this

The stressed syllables are “talked”, “voice” and “this”. Hearing the two prosodic gestalts, some might argue that this line consists of two intonation phrases. However, following the stressed syllable “talked” there is no slowing down of tempo or lengthening of syllables, typical features of unstressed syllables following the nucleus of a complete intonation phrase. Instead, the speaker speeds up the tempo, raises the pitch to the very top of her vocal range, and increases the volume.

The following wave form supports the auditory perception:

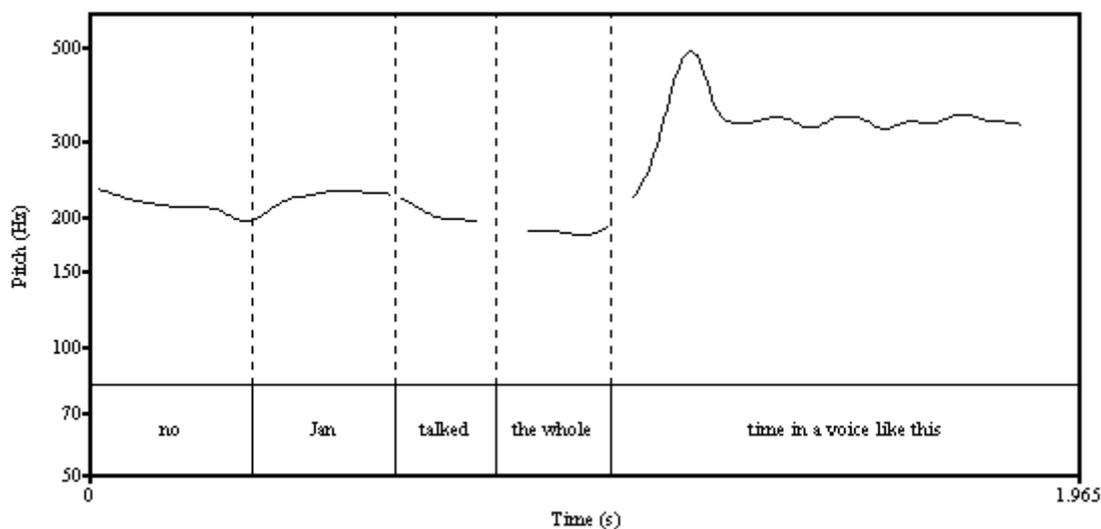
3.6



The wave form shows that the intensity suddenly increases starting with the syllables “a voice.” Furthermore, starting with “time”, the syllables come at a much quicker tempo.

An acoustic analysis of the fundamental frequency also reinforces the impression of two prosodic gestalts within one intonation phrase:

3.7



The interactants orient towards the nested prosodic gestalt first through laughter: Alina begins to laugh in line 3. L's laugh particles are found throughout the rest of the sequence. Not only laughter shows orientation towards the prosodic design of the gestalt, Alina continues the next intonation phrase using the prosodic design started in line 2, speaking loudly, in a very high register, and in a fast pitch. As in the example "Mad" above, she stylizes the gestalt even more, speaking with a higher pitch and a higher volume than in line 2.

4 Prosodic Gestalts above the Level of the Intonation Phrase

Not only are there gestalts at or below the level of the intonation phrase which are interactionally relevant, as shown above, there are also gestalts larger than the intonation phrase. Moreover, such gestalts may also be collaboratively created by more than one interactant.

Although collaborative completions have been the topic of research from an interactional point of view (cf. Ferrara: 1992; Lerner: 1991, 1996 *et al.*), few have considered the prosodic dimension. Beatrice Szczepek (1998, 2000) is one

exception. She has found that prosodic integration or non-integration of a completed second is important for determining whether or not the content of the utterance is attributed to the first speaker, the second, or to both. According to Szczeppek “most prosodic completions are integrated into the prior contour (...). Yet there are some that are prosodically prominent (...). The fact that they are prosodically prominent does not necessarily make them independent of what has gone before. They are still integrated into the preceding contour” (2000: 18).

When speakers collaboratively create complex gestalts, they display a heightened orientation to the prosody of the other speaker. Such gestalts may be quite intricate, and cover a larger stretch of the speech continuum than the intonation phrase does. Consider, for example, the following excerpt:

CUZ: 7:45 “Ugliest Set of Shoes”

In the same conversation as above (“Ugliest set of shoes”, “In a voice like this”) Lenore tells her cousin Alina about meeting another relative, this time her brother-in-law, Antonio.

1 A: Antonio had on white shoes
2 L: [I'm sorry
3 A: [I don't care if he is Italian
4 L: [ha
5 A: [I don't care
6 that was the ugliest set of shoes I ever saw in=
=my life
7 ha ha ha ha
8 L: .hhh
9 A: one of the kids at work said
10 → but Alina it doesn't count unless he had a white=
11 → =belt
12 → <clap clap>
13 → and he did
14 → L: he did ha
15 → A,L: ha ha ha ha ha ha
16 A: I [don't care if it's a white suit
17 L: [ha
18 A: put something else ha[ha ha
19 L: [ha ha ha

Above in section 2 it was argued that a stylized auditory gestalt below the intonation phrase occurs in line 6. A little bit later, however, a stylized auditory gestalt above the intonation phrase also occurs. The evidence of its existence can be seen through the orientation towards it by the second speaker, L, who continues the gestalt in a collaborative way.

In this excerpt, Alina is talking about her meeting with her brother-in-law, Antonio, and is discussing his bad taste in clothes. After asserting that his shoes were “the ugliest” she had ever seen in her life, she introduces direct speech in line 9: “one of the kids at work said”, and then she presents the direct speech in line 10: “but Alina it doesn’t count unless he had a white belt”.

Alina takes great care in developing the unfolding narrative, whose point is to show Antonio’s extremely bad taste in clothes. Already her opening, “Antonio had on white shoes” gets the response “I’m sorry” from Lenore. After L’s apology for Antonio’s bad taste, Alina sets up her first punch line: “I don’t care if he is Italian”, followed by a single laugh particle by Lenore. Alina then states again “I don’t care”. The restart is designed to highlight the first punch line, which, as shown earlier, is highly stylized: “that was the u::gliest set of shoes I ever saw I my life”. This is followed by Alina’s own laughter and audible out-breathing by Lenore.

Alina then proceeds to set up the next punch line. She begins by stating what “one of the kids said at work said”, which she gives in direct speech: “but Alina it doesn’t count unless he had a white belt” (line10-11). By designing her turn in this way, i.e. by putting the utterance in the mouth of a third party who is not present, the judgment of Antonio’s bad taste in clothes is not only Alina’s own subjective opinion but is heard rather as more universal.

I argue that the single gestalt now collaboratively produced by both speakers allows them to share the punch line of the developing joke as well as the laughter following it.

An isochronous rhythmic pattern marks the speech which follows the passage in direct speech. First, Alina claps twice, and then gives the punch line, “and he did”.

Lenore repeats word for word “he did”, thereby recycling Alina’s words. Alina’s contribution is also rhythmically integrated in the unfolding isochronous rhythmic gestalt. Her repetition of the punch line is then followed by rhythmic laughter by both participants. The rhythmic laugh particles are also integrated into the isochronous gestalt.

The rhythmic gestalt can be shown in the schematic representation below:

```
1 /<clap> /  
2 /<clap> /  
3 /and he /  
4 /did he /  
5 /did ha /  
6 /ha ha /  
7 /ha ha /  
8 /ha ha
```

Not only is this excerpt isochronous, following the clapping, it is also isometric in Frank Müller’s sense (1989, 1991). Between each stressed syllable which falls at equal points in time, there is the same number of syllables. In this example, each measure contains two beats: one strong beat and one weak beat. Even the laughter can be heard as maintaining this strong/weak pattern. Not only is the isochronous rhythm collaboratively created, so is the isometric rhythm. In line 4, Alina’s strong beat “did” is followed by Lenore’s weak beat “he”. In line 5, Lenore’s strong beat “did” is followed by Alina’s weak laugh particle. In lines 6-8, the laugh particles can be heard as being strong and weak, thereby continuing the isometric and isochronous pattern set up earlier.

Current prosodic research would consider the above excerpt as having two intonation phrases: Alina’s “and he did” and Lenore’s “he did”. The clapping and laugh particles, although they would be transcribed, would not be an integral part of the gestalt. The isochronous and isometric rhythm, started by Alina with her clapping and ending with rhythmic laughter, marks a gestalt which is larger than the intonation phrase and shows, through its finely collaborative character, its interactional relevance.

5 Summary

The preceding chapter shows that interactants orient towards intricate prosodic structures at, below, and above the level of intonation unit. Following Hutchby and Woofitt (1998) and Ono and Thompson (1995), I show how the display of orientation towards a particular phenomenon (in this case compound, nested and complex gestalts at, below and above the level of the intonation unit) indicates the interactional relevance of the phenomenon for the participants themselves.

Whereas displays of orientation give evidence for the proposed prosodic structures, the systematic function of these structures within certain discourse activities must be explored further. I propose that interactants form their contributions in prosodically intricate ways in order to guide participants understanding of the unfolding activity. Further research which specifically looks at compound, nested, and complex gestalts within certain activity types could give a greater insight into how these structures, used as practices, actually function.

Chapter 3: A Gestalt Approach to the Highlighting and Stylization of Prosodic Structures

The following chapter intends to describe the processes with which, in everyday spoken discourse, interactants manipulate the prosody of their spoken contributions into “highlighted” and “stylized” gestalts. The term “highlighting” is introduced to describe how gestalt elements may be manipulated in order to draw attention to an individual (i.e. “local”) part of a gestalt. Within spoken interaction such manipulation often coincides with an increased orientation to the stylized structure by conversation participants. The term “stylize” is introduced to describe how interactants manipulate gestalt elements and structures in order to bring out an underlying structure of an entire gestalt. In contrast to highlighted gestalts, stylized gestalts are, by their very nature, more global. Like highlighted gestalts, such stylized prosodic gestalts also elicit increased orientation to themselves.

1 Introduction

After close analysis of data from our corpus of spoken discourse, I assert that gestalt highlighting and gestalt stylization consist of three processes: 1) the “reduction” of irregular individual parts of the gestalt, 2) the “adjustment” of irregular individual parts into regular individual parts, and 3) the “ornamentation” of certain individual parts by overlaying additional gestalt elements on top of an underlying gestalt. The three gestalt processes make different hierarchical structures stylized: reduction reduces or suppresses local variations, thereby bringing out more global structures that are hidden or obscured by these local variations; ornamentation brings out local structures by highlighting them; and adjustment draws attention to both local and global structures. Highlighted and stylized gestalts may combine a variety of reduction, adjustment and ornamentation processes in order to produce a prosodic structure which is particularly salient.

To demonstrate the processes of “reduction”, “adjustment” and “ornamentation”, I will first present visual examples based on Max Wertheimer’s concept of *Prägnanz* (1923, 1938). I will then show how these processes can be applied to temporal, auditory gestalts in music. Musical examples are based on Lerdahl and Jackendoff from their work “A Generative Theory of Tonal Music” (1983). Finally, to demonstrate that these processes are also exploited by interactants in spoken discourse, I will present examples from our corpus of spoken discourse.

In order to underpin what one perceives auditorily in conversation, acoustic analyses – performed with the freeware program PRAAT²⁶ – will be presented. Such acoustic analyses will be used to show, for example, how in a segment of stylized speech specific frequencies may repeatedly occur; furthermore, these segments may even be lengthened and manipulated so that harmonic intervals, measured in semitones, are produced. Although an acoustic analysis does not in every instance match auditory perception (a computer lacks, for example, the cognitive processes that are integral for perception), it frequently does underscore auditory perception. For that reason, such analyses are presented as additional evidence that speech is indeed highlighted and stylized in the manner suggested.

“Highlighting” and “stylization” are important concepts not only in the fine arts, but, as I assert, in prosodic studies as well. The following section analyzes the terms “highlighting” and „stylization“ from both a gestalt theoretical perspective as well as from a social-linguistic perspective. The formal gestalt theoretical definition of “stylization” is combined with a pragmatic social-linguistic definition in order to provide a theoretical understanding of highlighting and stylization which is empirically grounded.

1.1 Gestalt Theory and Stylization

The fundamental terms “gestalt”, “gestalt elements”, and “individual gestalt parts” are taken over from gestalt theory as set down by its most renown proponents: Christian von Ehrenfels, Max Wertheimer, Wolfgang Köhler, Kurt Koffka as well as by later

²⁶ For information on the software program PRAAT see the following homepage: <http://fonsg3.hum.uva.nl/praat/praat.html> or contact Paul Boersma at boesma@fon.hum.uva.nl

researchers such as Wolfgang Metzger *et al.* Metzger in his essay “Grundbegriffe der Gestaltpsychologie” (1986), explains the two essential terms “gestalt” and “gestalt quality”:

Der Ausdruck „Gestalt“ wird in der Gestalttheorie, wie schon im alltäglichen Sprachgebrauch, in zwei Bedeutungen verwendet: einer substantiellen oder substantivischen, die auch in der Mehrzahl auftreten kann, und einer akzidentellen oder adjektivischen, die nur in der Einzahl auftritt: Substantivisch bedeutet Gestalt ein Ganzes von einer bestimmten Art; adjektivisch bedeutet es eine bestimmte Eigenschaft eben dieser Art von Ganzen. [...] In jedem Fall aber handelt es sich um Ganze, das heißt um überpunktuelle Gebilde oder Sachverhalte, die räumlich, zeitlich oder raumzeitlich ausgedehnt sind, mit Eigenschaften, die sich nicht aus artgleichen Eigenschaften der punktuellen Elemente herleiten lassen. Diese Eigenschaften nannte v. Ehrenfels „Gestaltqualitäten“ [...]. (Gestaltqualitäten) sind nämlich diejenigen Eigenschaften eines Ganzen, die *beim Herausblenden* immer kleinere [*sic*] Teilbereiche aus dem Ganzen (mit dem optischen Reduktionsschirm oder analogen Einrichtungen für andere Sinnesgebiete) [...] *verschwinden*.

(125)

In gestalt theory the expression “gestalt” has – as in everyday speech – two uses: a substantial or substantive use (which can also appear in the plural) and an accidental or adjectival use (which can only appear in the singular). In its substantial sense, gestalt means a whole of a particular kind; in its adjectival sense, it means a particular characteristic of this whole. [...] In any case it is concerned with the compositional whole, i.e. with structures or items that unfold in time or place and with characteristics which may not be deduced by similar kinds of characteristics of the individual elements. Christian von Ehrenfels calls these characteristics “gestalt qualities”. Gestalt qualities are those characteristics of a whole which disappear when the smaller individual parts of the whole are removed (with optical reduction blinds or similar apparatuses for other senses).

(my translation)

Metzger defines “gestalt” as a perceivable structure composed of individual parts but which is unlike any of the individual parts. The “individual parts” (“*punktuelle Elemente*”) contribute to the make-up of the entire gestalt but are fundamentally different from the entire gestalt structure. Furthermore, the entire gestalt has particular “gestalt qualities” (“*Gestaltqualitäten*”) that are not necessarily to be found in the particular characteristics of the individual parts of the gestalt. As Metzger

asserts, one cannot achieve the characteristic qualities of the gestalt through any of the characteristic qualities of the individual parts. The quality or characteristic of the entire gestalt is therefore fundamentally different from its individual parts.

The concepts of “highlighting” and “stylization”, as proposed in this dissertation, have been central to gestalt theory and go back to the earliest works in gestalt theory. One already finds reference to such processes in Christian von Ehrenfels’ groundbreaking essay from 1890 “On ‘Gestalt Qualities’” as well as in later essays, such as his 1916 “Gestalt Level and Gestalt Purity”²⁷. Von Ehrenfels describes the qualities of a highly stylized gestalt as “Gestalt qualities of a higher order”. In the latter essay von Ehrenfels introduces such concepts as “degree of formedness” (*Gestaltung*), “unity and multiplicity” (*Einheit* and *Mannigfaltigkeit*), “gestalt level”, and “gestalt purity”.

Von Ehrenfels asserts that every gestalt has a degree of formedness, which gives it a certain gestalt level (118). According to von Ehrenfels, “a rose is a gestalt of higher level than a heap of sand: this we recognize just as immediately as that red is a fuller, more lively colour than grey. The higher gestalts are distinguished further from the lower by the fact that the product of unity and multiplicity is in their case greater. For a fixed degree of multiplicity of parts, those gestalten are the higher which bind this multiplicity into a stronger unity” (118).

Ehrenfel’s concepts of “purity” and “gestalt level” (“*Steigerung*”) are analogous to the proposed processes “reduction” and “ornamentation”. According to von Ehrenfels, “a further characteristic of gestalten [...] is that of *purity*. This, too, is gradual in nature, but it is distinguished from gestalt level by the fact that it possesses an unsurpassable maximum, where the raising (*Steigerung*) of a gestalt is thinkable *ad infinitum*. The ideal forms of the mathematically exact sphere and of the regular polyhedra are gestalten of maximal purity, i.e. it is not even logically possible for this purity to be surpassed, but they are of a relatively low gestalt level” (119). According to von Ehrenfels, purer gestalts are therefore gestalts which can be *reduced* no

²⁷ Originally published in von Ehrenfels’ *Kosmogonie*, Jena: Diederichs, 1916, 93-96 1960: 44-46. Translated by Barry Smith in *Foundations of gestalt Theory*, Philosophia Verlag GmbH, 1988, 188-120).

further; that is, there can be no further removal of individual gestalt parts without fundamentally changing the gestalt. Gestalts of a high level are complex (what I term “ornamented”) gestalts which display both unity and multiplicity.

The proposed concepts of highlighting and stylization are based on formal gestalt theory. The following will now look more closely at the term “stylization”, particularly from an empirical “language-in-use” point of view.

1.2 Stylization: A Language-in-use Perspective

The terms “stylization” and “style” have been used in many contexts with different shades of meaning by researchers in a variety of fields, including, among others, literary science, rhetoric, sociology, linguistics, etc. Etymologically, as Frank Müller points out, the Latin term “stilus” refers to the metal writing utensil used for engraving letters into wax tablets (1989a:61). The stylus had a pointed end for engraving and a flat end for correcting mistakes by rubbing over the letters in the wax. The current meaning of style and stylization came about as the term became used in a metaphoric sense, based on the idea that a stylus can be used not only to write, but also to correct and improve a written text. According to H.U. Gumbrecht (1986):

Ohne Zweifel assoziierte man die buchstabenlöschende Wirkung der abgeplatteten Seite des “stilus” mit dem Tilgen überflüssiger Sprachelemente, und so wurde die “Stil”-Arbeit zu einem Synonym für die Suche nach einer Eleganz und Prägnanz der Schlichtheit, welche ihrerseits wohl im ersten vorchristlichen Jahrhundert mit dem Übergang von mündlicher zu schriftlicher Performanz assoziiert wurde.

Without a doubt one associates the effect of erasing the letters with the flat side of the “stilus” with the removal of redundant language elements. Consequently, “style” became a synonym for the search for elegance and stylization in simplicity, which in the first century BC was associated with the transition from spoken to written performance.

(my translation)

Today, the terms “style” and “stylization” are broad and diffuse, and they differ in meaning depending on the object of study. From a social-linguistic perspective,

Volker Hinnenkamp and Margret Selting's book *Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik*, attempts to describe the meaning and function of "style" and "stylization" in spoken communication. Speaking of the various contributions to their book, Hinnenkamp and Selting assert:

Was die Beiträge [...] demgegenüber vereint, ist ihr gemeinsames Verständnis, die aktive und sinnherstellende Bedeutung und Funktion von Stil und Stilisierung – vor allem in mündlicher Kommunikation – anhand der Analyse von Transkripten aus face-to-face Kommunikation [...] im Detail nachzuweisen und damit eine "interpretativ-soziolinguistische Stilistik/Stilanalyse" zu begründen.

(1989:1)

What unites the contributions is their similar notion of proving in detail the active construction of meaning and function of style and stylization – particularly in spoken communication – through the analyses of transcripts of face-to-face communication, thereby founding an interpretive social-linguistic stylistic/style analysis.

(my translation)

Leo Hickey, in his introduction to the monograph, "The Pragmatics of Style" (1989), has a similar goal as Hinnenkamp and Selting to apply the terms "style" and "stylistics" to language-in-use contexts with the purpose of founding a so-called "pragmastylistics" (1989:8). According to Hickey, pragmastylistics is "stylistics with a pragmatic component. (It) can be described as a study of language-in-use which pays special attention to the choices made from among the various grammatically correct ways of expressing one and the same thing, which is semantically or truth-conditionally equivalent" (1989:8).

Researchers analyzing language in use have variously defined "style" and "stylization" in a broad sense, and then applied the term to the specific object which they are analyzing. Broadly defined, ethnographer Dell Hymes sees "style" as "[...] a way or mode of doing something" (1981:34). Focusing more on the specific elements which make up "style", Peter Auer defines "Stil" as the

Menge interpretierter, kookkurierender sprachlicher und/oder nichtsprachlicher Merkmale, die (Gruppen/Rollen von) Personen, Textsorten, Medien, etc. zugeschrieben werden.

(1989:30)

set of interpreted co-occurring spoken and/or non-spoken features which are associated with (groups/roles of) people, text types, media, etc.

(1989:30)

Depending on the object of study or on one's research perspective, the term "style" or "stylization" is often preceded by a modifier which restricts the use of the term. Frank Müller, for example, speaks of "lautstilistische Muster" (1989a) and "Mündlichkeitsstil" (1989); Inken Keim and Johannes Schwitalla speak of "soziale Stile" (1989); Susanne Uhmann speaks of "Interviewstil" (1989) and "Kommunikations- oder Interaktions-Stil" (1989); Ingwer Paul speaks of "pastoraler Stil" and "Stilisierung des Rituals" (1989); Helga Kotthoff speaks of "konversationelle Stile" (1989); and Margret Selting speaks of "Gesprächsstile" and "Sprechstile" (1989, 1997), etc. The various modifiers demonstrate the various acts and elements which can be "stylized" and give the misleading impression that there is a uniform meaning of the broader terms "Stil" and "Stilisierung".

Although the above named researchers use the same terms "style" and "stylization", the meaning that underlies these two terms varies greatly. Volker Hinnenkamp points this out when he notes that "Tatsächlich ist Stilisierung zunächst ziemlich weit entfernt von Stil" (*in fact stylization is initially rather far removed from style*) (1989:253) Whereas "to style something" is often used in a sociological sense (according to Hinnenkamp ("Stil [...] beinhaltet interaktionsstrukturelle, letztendlich linguistische Beschreibbarkeit" (*style implies the ability to be described through interactional structures and, ultimately, linguistically*)), "stylization" is often used to describe perceptible gestalts. Again, Hinnenkamp asserts:

Beim Stilisieren steht der gestalterische Effekt im Vordergrund: Es wird Gestalt gegeben, die Form vereinfacht – künstlerisch insofern, als Komplexität reduziert wird auf interpretierte Kernmerkmale, die aufgrund

ihrer semantischen Kraft eine ganzheitliche Subsumption des so gestalteten ermöglichen.

(1989:253)

With stylization the gestalt effect is foregrounded: something is given form (gestalt), the form is simplified – artistically insofar as complexity is reduced to a few interpretable core characteristics which, because of their semantic power, allow a whole subsumption of that which has been formed.

(my translation)

In the following chapter, it is not the sociological sense of “style” which I mean when I use the term “stylization”, but the gestalt sense of the term. “Highlighting” refers to the manipulation of an individual part of a gestalt at the local level, thereby drawing attention to this part of the pattern. “Stylization” refers to the manipulation of individual parts of an entire gestalt, thereby drawing attention to the underlying structure of the whole pattern.

The chapter is divided into five parts: the following section 2 deals with the concept of “highlighting” and demonstrates highlighting processes on individual gestalt parts. Sections 3, 4 and 5 deal with the concept of “stylization”. Each section demonstrates the highlighting processes of reduction, adjustment, and ornamentation using visual, musical, and prosodic examples. Finally, section 6 shows a combination of highlighting and stylization processes in the creation of highly complex, stylized prosodic gestalts.

2. The Highlighting of Individual Gestalt Parts

Within an entire gestalt structure, certain individual parts may display a higher gestalt level via the manipulation of certain parameters. Such manipulation may be the result of the addition of gestalt parameters, which I shall call “ornamentation”, or the elimination of gestalt parameters, which I shall call “reduction”.

For the highlighting of a gestalt part – visual, musical, or prosodic – the processes of ornamentation and reduction are exploited for drawing attention to a smaller segment of the larger gestalt structure. In the following sections, I will demonstrate gestalt highlighting with visual, musical and, finally, prosodic gestalts.

2.1 Highlighting Processes in Visual Gestalts

Consider the following visual gestalts, 2.1.1 and 2.1.2:

2.1.1 Visual Gestalt:

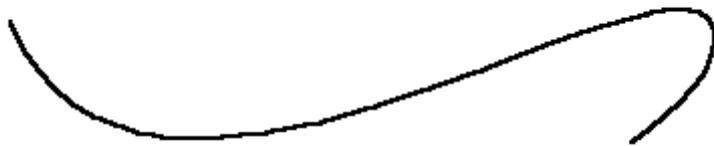


2.1.2 Reduced Highlighted Gestalt



In gestalt 2.1.1, the sweeping line displays local irregularities. At a particular part of gestalt 2.1.2 – the “hook” on the right side of the swoop – these local irregularities are suppressed, that is, local irregularities are removed, resulting in a finer (in von Ehrenfel’s terms a “purer”) line. Further reduction of these local variations across the entire gestalt structure would produce the following stylized gestalt 2.1.3:

2.1.3 Stylized Visual Gestalt



Whereas reduction may be thought of as a highlighting process that removes irregular local variation, ornamentation is a highlighting process which adds gestalt parameters on top of part of the gestalt structure. Consider the following gestalt, 2.1.4:

2.1.4 Highlighted Visual Gestalt



The difference between visual gestalts 2.1.3 and 2.1.4 is the changing of gestalt parameters through the addition of another gestalt parameter: color. The addition of red draws attention to this (local) part of the gestalt, therefore making it more stylized. The process of adding a gestalt quality on top of a gestalt structure I term “ornamentation”.

The following section, 2.2, will now demonstrate the highlighting of individual gestalt parts using temporal, auditory examples from music.

2.2 Highlighting Processes of Reduction and Ornamentation in Musical Gestalts

Not only in visual gestalts, but also in musical gestalts one can observe the highlighting of individual gestalt parts. The following will demonstrate the highlighting processes of reduction and ornamentation on individual parts of a musical gestalt.

To demonstrate reduction and ornamentation in a musical context, consider the following Irish jig, “Tripping up the Stairs” (also called “Sackow’s Jig”):

2.2.1

Tripping Up The Stairs



Typically in Irish music, a jig (a dance tune in 6/8) consists of two parts, A and B. Each part consists of eight bars which are repeated in the form AA BB. To bring out an individual part of the gestalt of the jig, a musician may manipulate or “highlight” various gestalt parameters of part of the musical segment. Although any part of the jig may be ornamented, for demonstration purposes, I will show ornamentation only on the first four bars. In these measures, the main melody of the A section is presented:

2.2.2



Using the gestalt element of volume, a musician might choose to highlight part of the original gestalt as follows:

2.2.3



In this gestalt, the first note is given extra volume. The result is an ornamented gestalt in which a certain local structural element, in this case the first note, is brought to the foreground and made salient. The process of ornamentation adds depth to the original gestalt structure, i.e. the note that is ornamented carries more weight than the other non-ornamented notes.

Another way of highlighting the underlying gestalt through ornamenting is through the addition of notes which harmonize with the original notes, such as in the following:

2.2.4



Collaboratively, musicians will create highlighted gestalts through the process of ornamentation. For example, if a musician playing the traditional Irish frame drum called a bodhran were to tap the drum at the first beat, thereby drawing attention to this part of the musical gestalt, then one could also speak of ornamentation in an highlighted gestalt:

2.2.5



Not only can a particular part of a musical gestalt be ornamented, as displayed above, it can also be reduced. Consider again the first four bars of the tune, “Tripping Up The Stairs”:

2.2.6



An Irish musician might stylize part of the gestalt by reducing the first three notes of the gestalt as follows:

2.2.7



Just as particular parts of musical gestalts can be highlighted through the processes of reduction and ornamentation, so can prosodic gestalts, as will be demonstrated in the following section. The above musical examples have shown how a particular part of a pattern may be brought out through the processes of ornamentation and reduction. An individual note within a gestalt, for example, can be made salient through the overlaying of additional gestalt parameters, such as extra volume, for

example. An individual part of a gestalt may also be made more stylized through reducing the notes to a single note, for example. It is proposed that such ornamentation and reduction processes, in which the gestalt is highlighted so that a particular part of a pattern comes to the fore, have interactive contextualizing functions. One such function of highlighting this particular first note might be as a signal to dancers that the dance has begun, for example. Or the highlighted note could be used to show that the tune is being repeated. Just as in spoken interaction, gestalt highlighting in music may have contextualizing functions²⁸.

2.3 Highlighting Processes in Prosodic Gestalts

Ornamentation in temporal auditory prosodic gestalts functions similarly as in musical gestalts. Interactants ornament certain parts of the prosodic gestalt in order to bring them to the fore. Often, the highlighting process of ornamentation coincides with the stylization process of reduction. Local irregular variations in the prosodic gestalt are suppressed, thereby bringing out important aspects of the more global prosodic gestalt. After the global gestalt is reduced, local parts of the gestalt may then be highlighted. To show how a prosodic gestalt may be highlighted through ornamentation, it is necessary to see how the underlying gestalt of an intonation contour looks in non-highlighted discourse. For this purpose, I shall use David Crystal's (1969) and Allen Cruttenden's (1983) descriptions of the intonation unit.

In his book, Intonation, Alan Cruttenden sets out to describe the various features of the prosodic system of English in spoken discourse, a task undertaken 13 years before by fellow linguist David Crystal. After asserting the "intonation group" as the basic unit of spoken discourse (1969: 25-45), Cruttenden proceeds to describe the various features which make up the intonation group. According to Cruttenden, "within each intonation-group we have a number of PITCH ACCENTS (which indicate prominent syllables and hence prominent words); among the pitch accents in an intonation-group, one is considered the NUCLEUS (indicating the most prominent syllable and hence most prominent word)" (1986:55). Stressed syllables include not

²⁸ Without pursuing the idea further here, I assert that some of the methods of conversation analysis could be fruitfully applied to the analysis of the interaction between musicians or between musicians and audience in a musical performance – particularly in informal settings such as in sessions.

only pitch accents, but also any prominent syllable that is brought to the foreground by virtue of lengthening or loudness.

According to Cruttenden, stressed syllables may have primary, secondary and tertiary stress. Pitch accent (that is, prominence due to pitch obtrusion), lengthening, and loudness are the features which, according to Cruttenden, differentiate stressed syllables into primary, secondary and tertiary stresses. Primary stress will fall on the nucleus and be indicated by pitch accent; secondary stress will fall on another prominent syllable and be indicated by pitch accent, lengthening, or loudness. Tertiary stress will be indicated by length and/or loudness.

For the highlighting of prosodic gestalts on the local level, speakers in discourse focus on one stressed syllable, which is usually, but not necessarily, the nucleus.

Nuclear pitch accent is an integral part of the intonation unit. In his taxonomy of English nuclear tones, Allen Cruttenden distinguishes between accents with high fall and accents with low fall, and accents with a high rise and accents with a low rise, such as in the following examples, taken from Cruttenden, 1983:59:

The image displays four examples of pitch accents for the sentence "Are you going a way?". Each example consists of the sentence written above a horizontal line, with a series of dots representing pitch levels and a final curve indicating the intonation contour.

- Example 1: "Are you going a[`]way?" - The pitch is high on the first syllable, falls to a low level on the second syllable, and remains low. The curve is a high fall.
- Example 2: "Are you going a\way?" - The pitch is high on the first syllable, falls to a low level on the second syllable, and remains low. The curve is a low fall.
- Example 3: "Are you going a/way?" - The pitch is high on the first syllable, falls to a low level on the second syllable, and then rises to a high level on the third syllable. The curve is a high rise.
- Example 4: "Are you going a'way?" - The pitch is high on the first syllable, falls to a low level on the second syllable, and then rises to a high level on the third syllable. The curve is a low rise.

According to Cruttenden, nuclear pitch accents can carry certain “meanings”:

The high-falling and low-falling tones both have rather “serious” overtones, the higher tone being more “involved” and the lower tone more “business-like”; whereas the rising tones are altogether “lighter” tones, the high-rising tone being the most casual.

(1986:59).

Unlike Cruttenden’s approach, which assumes contextual meaning on constructed examples, the following will show gestalt highlighting using empirical data. The examples will show some of the various ways speakers may ornament particular parts of the prosodic gestalt through the manipulation of gestalt elements on particular pitch accents within the intonation contour. Furthermore, CA methods will be applied to show how such highlighting of the prosodic gestalt is interactionally relevant.

Consider the following example:

2.3.1 “Chicken People”

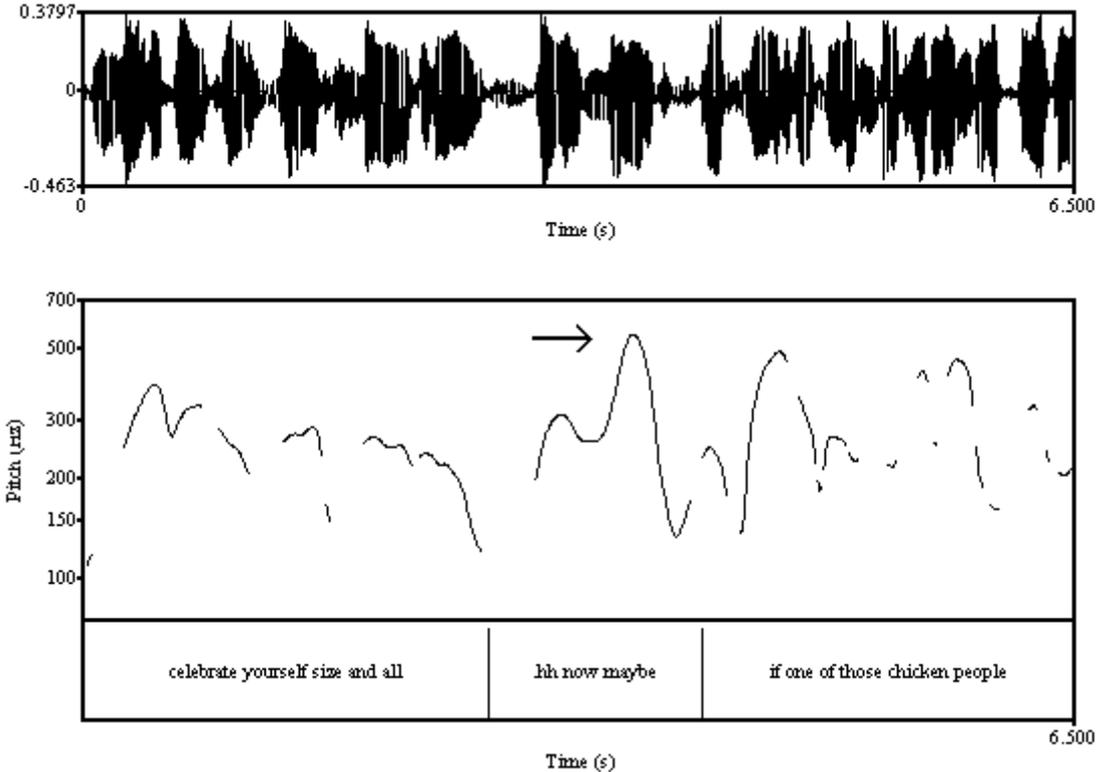
In her radio talk show program, “The Barbara Carlson and Friends Show”, Barbara Carlson (B) jokes with her fellow moderator, Peter Theo (P), about mayoral candidates of Minneapolis and St. Paul recognizing the rights of overweight and ugly people.

```
1 B: pInk letters
2   on a chOcolate frOsted cAke
3   .hh spElled out the mEessage
4   cElebrate yourSElf sIze and All
5   → .hh now <<falsetto> ↑MAY>be
6   if one [of those (.) chIcken (.) pEople
7   P:      [.hh ha   ha ha ha ha   ha ha ha
8   B: that are [rUnning for mAyor of mInneapolis or st pAul
9   P:      [ha ha ha
10  B: could gEt themselves out above the crOwd
11     by suggEsting that we should accEpt ourselves
12     sIze and All
```

The above excerpt is taken from a section in which Barbara Carlson and her fellow moderator Peter Theo discuss in a playful way the rights of large and ugly people. In

lines 1 through 4 of this excerpt, Carlson starts by describing a cake with a frosted message that tells large people to celebrate themselves, including their being overweight. Following line four is audible in-breathing and the next utterance: “now maybe/if one of those chicken people”. Here, the second syllable “may” displays extreme pitch movement, not realized in any of pitch accents before. The highlighting on this pitch accent can be clearly seen in the following frequency analysis:

2.3.2



On the nuclear syllable of the intonation phrase, Carlson’s pitch jumps up to an exceptional 500 Hertz and then falls to about 200 Hertz. The top of the jump is indicated in the frequency analysis by an arrow. Not only is this single syllable highlighted with high pitch, but Carlson additionally highlights it with what Laver terms “ventricular falsetto”, which is produced “by very severe compressive effort of the whole larynx, and extreme pulmonic effort” (1980:139). The intensity of the syllable, as indicated by the wave form, is not exceptionally high. Due to the extreme compression on the larynx, however, Carlson signals loudness without actually being loud. Laver asserts that this voice quality “seems never to be used in normal speech”

(139). The combination of extreme pitch and unusual voice quality on this syllable highlights it and separates it out from the speech continuum.

Following the extreme pitch movement in line 5, Peter, in overlap with Carlson's next contribution, begins to laugh (line 7). His laughter at exactly this point can be taken as an indication that he is orienting towards Carlson's prosodic highlighting.

The important question, why does Carlson highlight this particular syllable the way she does, can be answered by looking at what is happening interactionally. In a playful modality, Carlson is criticizing politicians by asserting that they do not address the rights of obese people. She sets up her complaint in lines 1-4 through the description of a cake with the decorated message "celebrate yourself size and all" written in pink frosting. The extreme pitch movement and volume on the word "MAybe" introduces the transition to her complaint about politicians not taking the rights of large people seriously. The extra volume and pitch movement cue the playful modality and are followed by a syntactically incomplete conditional: "if one of those chicken people who are running for mayor of Minneapolis or St. Paul could get themselves out above the crowd by suggesting that we should accept ourselves size and all." The gestalt pattern is abrupt and startling, and strategically helps get her point across in an emphatic way. Additional evidence for Carlson's successful strategy is shown when mayoral candidate John Derris calls her program and asserts that "I called to tell you that in a shameless effort to pander for your support to get that lawn sign that immediately upon being elected mayor of Minneapolis I am going to declare size acceptance month throughout the city of Minneapolis and not only that but the bigger you are the more we'll accept you."

Whereas Peter in the above example displays his orientation towards the highlighting of the prosodic gestalt through laughter, in the following the recipient displays her orientation towards the prosodic gestalt through repetition. The excerpt is taken from a telephone conversation between two friends, B and F:

2.3.3 “Larry’s Story”

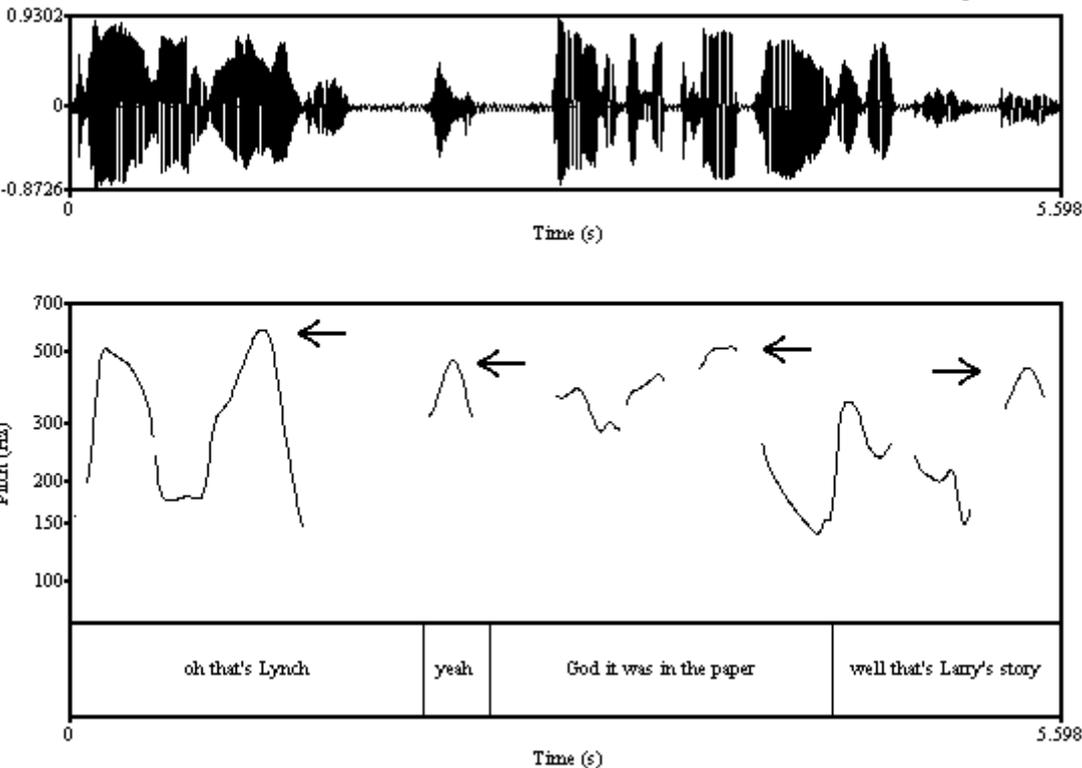
After first apologizing to F for keeping her power tool too long, and then thanking her for a dinner which F organized, B tells F how much she respects F and F’s husband’s work. In the following excerpt, one of F’s husband’s projects appeared in the newspaper, and B realizes the connection between the newspaper article which she read, and F’s husband.

1 B: [fantastic
2 F: [???
3 B: he [can’t miss
4 F: [but uh
5 B: you both can’t miss
6 F: that’s a different deal
7 so uh
8 he’s on (.) for them too
9 B: um hmm
10 F: and then he has a this a
11 College Park
12 which is up by FlOyd [???
13 B: [oh that’s L↑↓Ynch
14 F: ↑↓yEAh
15 B: it was in the ↑pA↓per
16 F: well that’s Larry’s ↑stO↓ry

Following each of B’s positive assessments of F and F’s husband, F does not directly acknowledge the compliments via “thank you”. Rather she attempts to refocus attention on her husband’s work, thus allowing her to negotiate a delicate situation. By responding with an informing following a compliment, F is able to show that she has heard the compliments, and it allows her to avoid agreeing or disagreeing, which are the kinds of second responses typical in assessment sequences. F’s informing about her husband’s work in lines 6-12 is responded to by B with a news receipt “oh that’s Lynch”. “Oh” here is heard as a change of state particle (Heritage 1984) and is prosodically highlighted (see chapter 4, section 2: “State Changes in Informing and Repair Sequences”, p.145). The syllable “Lynch” rises to a very high frequency of 600 Hertz and then falls abruptly. That this prosodic highlighting is heard and oriented towards is seen in lines 14, 15, and 16, in which a similar high rise/fall is heard. F repeats the prosodic highlighting in 14 with her utterance “yeah”, which is also realized very high in her voice range: approx. 500 Hertz. The prosodic highlighting is taken up again in line 15, when the pitch of the first syllable of “paper” rises again to approximately 600 Hertz, and falls sharply on the second syllable. Finally, F repeats the highlighting again in line 16 with a high rise on the first syllable

of “story” followed by a sharp fall on the second syllable. That both F and B repeat the prosodic highlighting shows that there is an orientation towards the highlighting of an individual part of a prosodic gestalt. Moreover, the repetition of the highlighted pattern creates a larger stylized gestalt which groups the speech continuum from line 13 though line 16²⁹.

The following frequency analysis supports the perception of the similarity of prosodic highlighting in lines 13, 14, 15 and 16:



2.3.4

The arrows point to the peak of the rise/fall pitch movement which is heard on “lynch”, “yeah”, “paper” and “story”. The repetition of the prosodic highlighting is regarded as proof that the highlighting of an individual part of a prosodic gestalt is interactively relevant to conversation participants.

²⁹ For a more detailed discussion of how repeated local variations bring out a larger gestalt pattern see p. 134.

Again, the question why do participants highlight their contributions can be answered interactionally. B, after complimenting F and her husband on their various accomplishments, discovers another accomplishment which she knew of, but had not associated with F and her husband. She contextualizes her change of state with the news receipt in line 13. F finds herself in the potentially embarrassing position of acknowledging yet another compliment after extreme complimenting in lines 1, 3, and 5: “fantastic”, “he can’t miss”, “you both can’t miss”. F acknowledges the accomplishment semantically in lines 14 and 16: “yeah”, “well that’s Larry’s story”, but her copying the prosodic highlighting of B (which contextualizes astonishment) allows F to display recognition and acceptance of the compliment without appearing arrogant. Through gestalt highlighting she successfully masters a potentially uncomfortable social situation: recognizing and accepting a compliment.

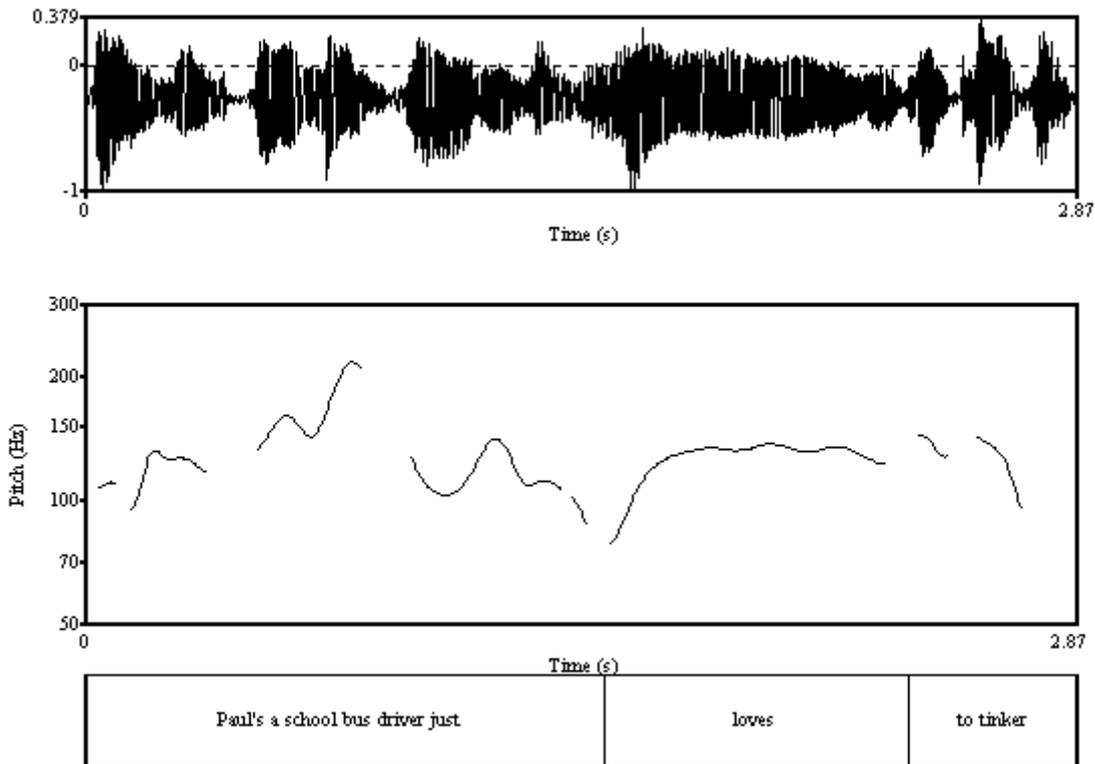
The above two examples showed that prosodic highlighting is oriented towards by conversation participants. The following examples will show other ways of manipulating the gestalt features of part of an individual gestalt for the purpose of highlighting. For an example of prosodic highlighting through local ornamentation via lengthened syllables, consider the following example:

2.3.5 “Loves to Tinker”

Paul Harvey (P), in his radio program which mixes news and advertisements, presents a news item in which a homemade model airplane caused confusion in local air traffic control.

1 P: Paul is confounding control towers (.) in Alaska
2 with a homebuilt airplane
3 that jumps off the ground within 75 feet
4 and it can be up at 2000 feet within 60 seconds
5 ATC at Juno
6 heard from an airline captain in the area
7 who said
8 what was that

9 ha ha ha
 10 Paul's a school bus driver just lo:ves to tinker



In line 10, Paul Harvey not only lengthens the syllable, “lo::ves”, he also keeps the tone at the same pitch of approx. 140 Hertz. This can be seen in the above frequency analysis, in which the frequency line flattens out above the syllable, “loves”, at a midpoint in Harvey’s voice range. Auditorily, such extreme syllable lengthening and tone manipulation is heard as “singing”.

In line 10 Harvey offers an explanation why Paul continues pursuit of his hobby although the results are potentially dangerous. The dangerous situations are a result of Paul’s love of tinkering rather than ill intent. The gestalt highlighting through syllable lengthening and held pitch has an iconic aspect. The drawn out singing of the syllable contextualizes the pleasure which Paul associates with his tinkering.

Besides extreme pitch movement and lengthening with pitch manipulation, a speaker may also ornament part of the gestalt through para-linguistic features, such as creak, falsetto, laughter, cracking voice, vibrato, etc. For a number of highlighting strategies involving the ornamentation of particular parts of a gestalt through the overlay of

prosodic and para-linguistic features, consider the following example, taken from a Christian revival worship service:

2.3.6 “I’m a Going to Die”

In an outdoor Christian revival worship service, the preacher gives testimony about how God has taken care of him.

1 I don't want to take up anymore of your time here tonight
2 we have other preachers n
3 I didn't feel like being up here
4 to begin with
5 but as I said I'm glad to be out there with you
6 one more time may be my last time
7 I don't know about <<crack, falsetto> that>
8 but let me
9 knowing that
10 knowing that I'm a going to (.) die:
11 brother Harry
12 knowing that God's a gonna call me home some of these-
13 [days
14 [clap clap
15 call me home
16 I call it brother Steve
17 and it is
18 my eternal home
19 I got a little home out yonder
20 waa I just gotta little home out there
21 God has leased me or let me live out there on a little-
22 piece of land brother Haggard
23 don't belong to me
24 truth of the matter it belongs to God
25 bless his name tonight it don't belong to me
26 I'm just he's he's jus he's just rented it to me
27 just let me live out there and jes given me
28 a givin me me time
29 ah ah ah brother he gives ah
30 he gives every man and every woman a chance ah
31 ah to seek him
32 every sinner man and sinner woman ah
33 a chance to seek him before he calls him home ah
34 I loves his name to let then lets praise his holy name ah
35 I want to ask the children of that are here to=
 <<vibrato> =night ah>
36 I loves his name tonight when you get on the firing=
 <<vibrato line> <clap>
37 and will you talk to men and women as you come in=
 =contact with them and some save
38 will you invite them to come on ah
39 or ask them to come on
40 come to the church ah
41 ah bless your holy name ah
42 ah ah ah bless your name [tonight
 [knock knock knock knock

42 a church
 43 of the church is a home for the children of
 44 yes it is it's a home
 45 it's a home for the children of
 46 and if there's any children
 47 if there's any Christian people in the house tonight
 48 any men and women in this house tonight that's been born=
 =again
 49 put that living water
 50 aw you're away from home
 51 I know you're not satisfied

The highlighting of parts of individual gestalts starts in line 7. In this line, the speaker's voice cracks into falsetto at the last syllable of the intonation phrase:

7 I don't know about <<*crack, falsetto*> that>

One indication that this is a highlighted gestalt rather than mere coincidence comes in the following lines, in which the last syllable of the intonation phrase becomes repeatedly the site for gestalt highlighting. In line 10, for example, the last syllable is followed by a short pause, lower volume, and syllable lengthening:

10 knowing that I'm a going to (.) die:

This part of the gestalt is again highlighted in lines 12 and 39, when the speaker overlays clapping and knocking at this point of the gestalt structure:

knowing that God's a gonna call me home some of=
 =these [days
 [clap clap

39 ah ah ah bless your name [tonight
 [knock knock knock knock

Further proof that the highlighting of this individual part of the gestalt is relevant is the speaker's addition of the semantically empty syllable "ah" which is heard at the end of lines 26, 27, 29, 30, 31, 32, 35, 37 and 38. This is usually realized as a lengthened syllable.

Finally, in lines 32 and 33 the speaker ornaments this part of the individual gestalt with the para-linguistic feature of vibrato:

I want to ask the children of that are here to <*vibrato*< night
ah>>
I loves his name tonight when you get on the firing <<*vibrato*
line>> <*clap*>

Through the repeated highlighting of the last syllable of the gestalt with the overlay of additional prosodic or para-linguistic features, one sees that the ornamentation of local gestalt structures is a relevant feature of discourse organization.

2.4 Summary

The above examples show how an individual part of a gestalt can be highlighted through the gestalt processes of reduction and ornamentation. Highlighted visual, musical, and prosodic gestalts may look as follows:

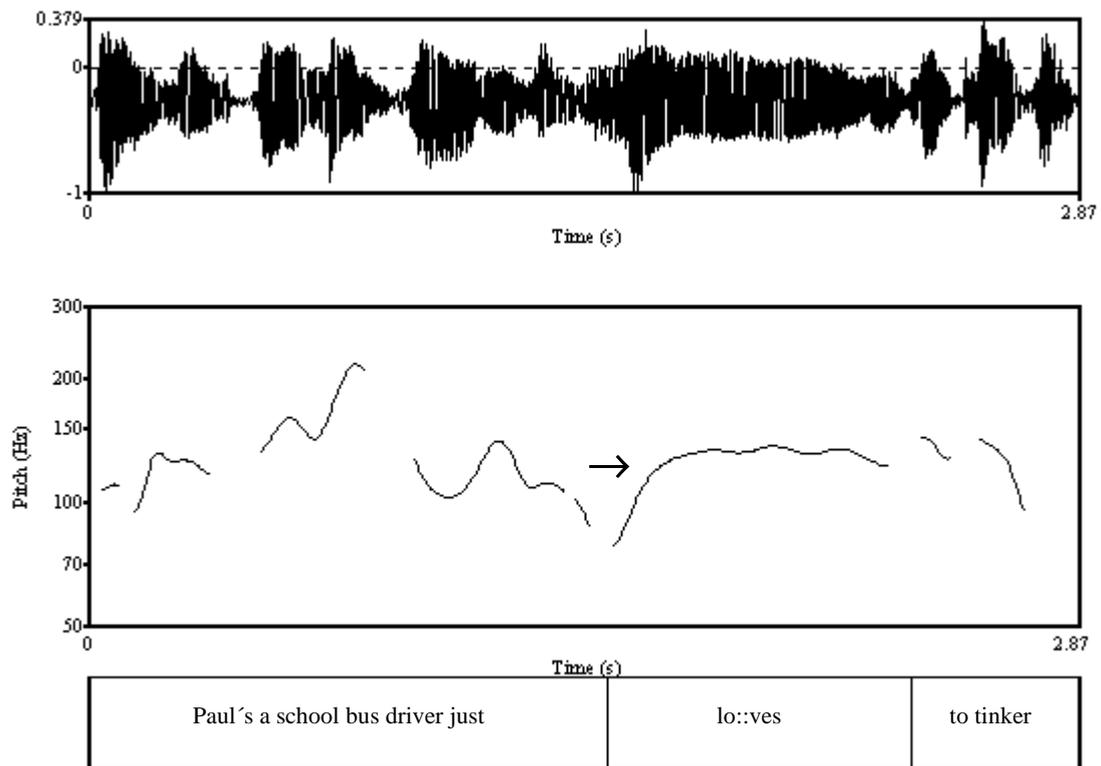
2.4.1 Visual:



2.4.2 Musical:



2.4.3 Prosodic:



In the visual gestalt above, highlighting is achieved through the overlay of color; in the musical gestalt, it is achieved through volume, and in the prosodic gestalt, through syllable lengthening. Furthermore, it has been shown that speakers in conversation, through laughter or repetition for example, orient towards the highlighting of individual parts of a gestalt.

The following will now turn to the stylization of gestalts.

3 Stylization in Visual, Musical and Prosodic Gestalts

In the above examples, individual parts of a gestalt were brought out through gestalt highlighting. There can also be a forming of the entire gestalt, however, in which case one can speak of “stylization”. For stylized gestalts, in which an entire gestalt structure is brought to the forefront, the processes of ornamentation, reduction and adjustment play an important role.

The process of reduction involves the suppression or deletion of local irregular variations over the entire gestalt. Such deletion or suppression results in the foregrounding of more global gestalt structures. For gestalts stylized via reduction, aspects of the original gestalt are removed or regularized, but the integrity of more global aspects of the original gestalt is maintained and made more pronounced.

Unlike the process of reduction, which removes irregular local variation, the process of “adjustment” converts irregular local variation into regular local variation. The result of regular local variation is a stylized gestalt.

To demonstrate gestalt reduction and adjustment, I will first consider visual, and then musical examples. Finally, I will show how these two gestalt processes are important aspects of spoken discourse.

3.1 Adjustment, Reduction, and Ornamentation in the Stylization of Visual Gestalts

In addition to the processes of reduction and ornamentation, another process for highlighting or stylizing a gestalt or gestalt structure is the process of “arrangement”. In adjusted gestalts local variations are not removed (as in reduced gestalts) but rather “adjusted” so that a regular pattern emerges.

Consider the following two visual gestalts, A and B, based on Max Wertheimer’s concept of good form (“*Prägnanz*” (1923)):

3.1.1 Unstylized Gestalt



3.1.2 Stylized Gestalt



In gestalt 3.1.2, local variations in the four angles are affected by the process of adjustment. Local variations are “smoothed out” or adjusted in such a way that a “*prägnante*” gestalt becomes apparent. According to Wertheimer, an observer of these two gestalts would perceive both gestalts as rectangles, but gestalt 3.1.2, with 90° angles at all four corners, would be perceived as more “*prägnant*” than the first. According to Wertheimer, “the observer frequently sees a right angle even when objectively a more acute or more obtuse angle is being presented, although the observer may report that it was “not quite correct”, “somehow wrong”, etc., he is usually unable to say in which direction the “error” lies (Wertheimer, 1923).

Gestalt B above is perceived as more “*prägnant*” than gestalt A. Both gestalts, however, are perceived as rectangles; the only difference in 3.1.2 is the suppression of local variation in the width of the angles. Through the process of reduction towards a more “*prägnante*” gestalt, gestalt 3.1.2 can be considered stylized.

The following examples also work with the process of reduction. Consider the following gestalts, 3.1.3 and 3.1.4:

3.1.3 Unstylized Visual Gestalt



3.1.4 Reduced Visual Gestalt



In the second gestalt, 3.1.4, the irregularities of the line which existed in the first gestalt, 3.1.3, are removed, reducing the gestalt to a more regular shape. Irregular local variations of the original gestalt 3.1.1 are suppressed, therefore allowing one to speak of reduction – and stylization.

Such gestalt reduction can also occur with other gestalt parameters, such as in the following visual gestalts 3.1.5 and 3.1.6:

3.1.5 Visual Gestalt



3.1.6 Reduced Visual Gestalt:



The multi-colored figure in 3.1.5 is reduced to black in 3.1.6. All gestalt parameters may be reduced. When a figure is reduced so that most of the original gestalt is still recognizable, then one can speak of stylization.

The following visual gestalt, 3.1.7, is a reduction of gestalt parameters in 3.1.4 above, but as the original gestalt is no longer recognizable, one cannot assert that it represents a stylized version of 3.1.4. In the following, not only are irregular local variations suppressed, but so are global variations.

3.1.7



The above examples have demonstrated how, on the local level, variations may be removed, resulting in a more stylized global gestalt. Such gestalt reduction as demonstrated in the non-temporal visual gestalts above may also be seen in musical contexts, as demonstrated in the following section.

3.2 Gestalt Reduction in Music

As in visual gestalts, stylization also occurs in musical gestalts when certain gestalt parameters are reduced, yet where the original gestalt may still be recognized. Lerdahl and Jackendoff have shown how, according to gestalt principles, a musical gestalt may be stylized by reducing the gestalt to its most structurally important elements (Lerdahl and Jackendoff 1983, 106 – 249). To demonstrate gestalt reduction in auditory temporal gestalts, I will use the following Irish fiddle tune, “Tripping Up the Stairs” (also called “Sackow’s Jig”).

Consider the following jig:

3.2.1



3.2.2



This first bar is unstylized in the sense that no additional gestalt manipulation is imposed on the gestalt. It is the basic gestalt which an Irish musician will then manipulate in order to bring out other gestalts. Unstylized, the first bar of “Tripping up the Stairs” consists of six eighth notes: f a a g b b. Hierarchically, the first and fourth notes of the gestalt are structurally more important in the underlying gestalt of a jig. According to Irish fiddle player Pete Cooper, “In a double jig [...] only the first of each group of three quavers (eight notes) is stressed” (Pete Cooper 1995:15).

The complex grouping of the musical gestalt is perceived according to gestalt principles. On the most global level, one perceives all six notes of the measure as its own gestalt, as demonstrated in the following:

(₁XXXXXX₁)

On a more local level, one perceives two gestalts nested within the more global gestalt. According to the gestalt principle of directionality, one hears the first three eighth notes as a single gestalt, and the last three eighth notes as a single gestalt:

(₁(₂XXX₂)(₃XXX₃)₁)

Finally, on the most local level, one may perceive each individual note as its own gestalt nested into the next higher global gestalt:

(₁(₂(₄X₄)(₅X₅)(₆X₆)₂)(₃(₇X₇)(₈X₈)(₉X₉)₃)₁)

A musician has many options to stylize the above gestalt. One option is to reduce the pitch of the original gestalt so as to bring out one of the more global gestalts in which the local gestalts are nested. By removing the second, third, fifth and sixth notes, one

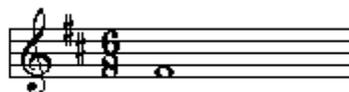
suppresses local variations and therefore draws attention to the following reduced gestalt, which is more stylized than the original:

3.2.3



Finally, these two gestalts can again be reduced to one single note, therefore bringing out the most global gestalt structure. The notes other than the first are suppressed:

3.2.3



In Irish music, gestalt reduction, as demonstrated above, has the function of drawing attention to the structure and of bringing out some of the nested patterns which are part of the original gestalt structure. An aesthetic experience occurs when a listener of Irish music who knows the original gestalt hears such a gestalt reduction.

It is believed that stylized prosodic gestalts function similarly. Such changes and manipulations in the original gestalt are not random, but follow gestalt principles. It should be pointed out that Christian von Ehrenfels, already in the late 19th century, observed that according to gestalt principles, although various gestalt parameters may change, the perception of the original gestalt may remain intact. Depending on the degree and type of gestalt reduction, however, the gestalt structure may become less and less familiar.

3.3 Gestalt Reduction in Prosodic Gestalts

Just as visual and musical gestalts may be stylized through the principle of reduction, so can prosodic gestalts. Unlike in the musical example above, in which the unmarked underlying gestalt is given, it is seldom that an original unmarked gestalt is first presented by an interactant and then immediately stylized in spoken discourse. In this sense, conversational discourse is somewhat less like Irish music and more like free jazz, in which the temporal auditory gestalts are spontaneously produced within the moment of improvisation. As improvisation exists on a spectrum from high to low in different genres of music as well as in speech, these differences may be seen as gradual and not fundamental.

This gradual difference between music and discourse, however, can be overcome. In order to see how an original gestalt is stylized, it is important to see what an unstylized prosodic gestalt looks like. Many linguists from a variety of linguistics schools – such as Crystal (1969), Cruttenden (1986), Laver (1994); Schuetze-Coburn/Shapely/Weber (1991), *et al.* – have used various methods to describe unmarked (i.e. unstylized) prosodic gestalts. The following will first present unstylized prosodic gestalts as described auditorily by Crystal, Cruttenden and acoustically by Schuetze-Coburn *et al.* After showing how unstylized prosodic gestalts appear in discourse, it can be seen how an underlying gestalt may be stylized through the process of gestalt reduction.

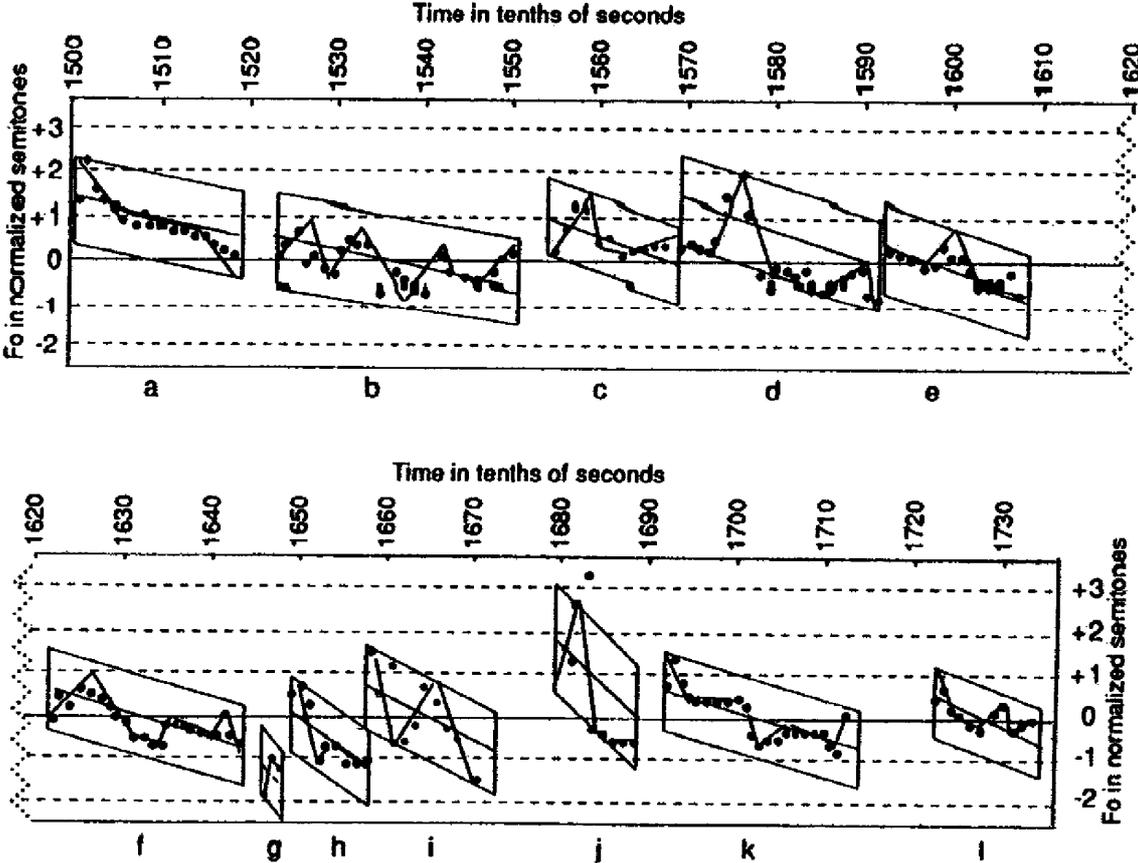
3.3.1 Gestalt Reduction in the Pitch of Prosodic Gestalts

Schuetze-Coburn/ Shapley/ Weber (1999) describe the intonation unit both acoustically and auditorily. According to their study “99% of the acoustic unit boundaries coincided with auditory unit boundaries” (1991:207).

In the following acoustic analysis, Schuetze-Coburn *et al.* (1991, 215) document the phenomenon of steadily declining pitch - measured here in semitones. Although there are irregular local variations in pitch, i.e. sometimes the pitch rises on the local

level, the general tendency on the more global gestalt level is for the pitch to fall. Following the acoustic analysis is the transcript of the speech utterances:

3.3.1.1



Schuetze-Coburn, et al. 1991:214-215)

3.3.1.2

- (1) (a) P: ... Uh nó I- I ... I gá- I méan
 (b) ... Í- I assumed- I was [relatively] cálm in the sence that I [[figured
 S: ... [[yes]]
 N: ... [[@@@@ (H)]]
 (c) C: After an hóur]] [and a hálf?
 (d) P: [Thát áfter a r- ... a réasonably shórt périod
 (e) I mean befóre the tóes begin to frée-ze
 X: @@@@
 (f) P: (H) That sómebody wóuld appéar ... I mean from ínsíde
 S: ... yeah
 (g) N: ... Mhm
 P: [I méan that-
 S: [[it's]]
 (h) N: [[(F Théy'd gét úp F)
 S: [[(he fígures he's]]
 (i) Léft a way óut ányway=
 ... of course
 (j) P: ... It wás the way óut
 (k) ... But [it túrned óut thére was anóther way out ... that I didn't
 (@ réalize @)
 X: ... [@@
 P: [[@@@
 S: [[oh
 (l) P: [(@ I shóuld've been «lóoking» óut @)
 X: [(P «mh» P)

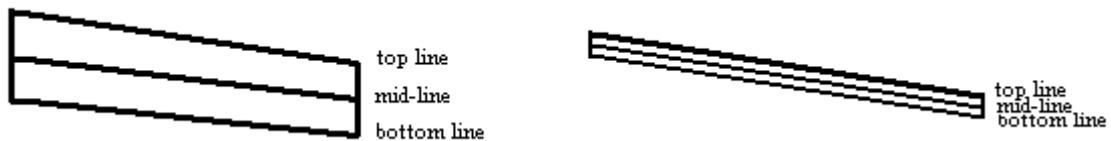
The above acoustic analysis describes twelve patterns, labeled by Schuetze-Coburn *et al.* a through l. Schuetze-Coburn *et al.* superimpose a parallelogram over the acoustic patterns of the resulting points of the frequency analysis. Each parallelogram defines a declination unit (= DU) consisting of three declining lines in which the bottom line connects the low points of the unit, the middle line connects the mid points, and the top line connects the upper points. "The end boundary of a DU was located where the F0 reached the speaker's lower bound, the bottom line of the parallelogram (the declination line) reached such a bound, or it was no longer possible to include high points in the envelope (parallelogram)" (213). This is what will be considered an unstylized prosodic gestalt in the following.

Through the process of reduction, an unstylized prosodic gestalt may become stylized. Interactants in conversation may suppress irregular variations, thereby bringing out a more global and, therefore, more stylized gestalt.

The pitch range, for example, may be compressed, thereby collapsing the wide parallelogram of an unstylized, unmarked gestalt from a top line, mid-line and bottom line into a narrow parallelogram in which the distance between the top line, mid-line and bottom line steadily decreases, as demonstrated in the following diagram:

3.3.1.3 Unstylized Pitch Gestalt

3.3.1.4 Stylized Pitch Gestalt



For an example of unstylized and stylized pitch gestalts, consider the following:

3.3.1.5 “Call the Man Stupid”

In this excerpt from the radio talkshow program “The Barbara Carlson and Friends Show”, the caller, David (D), complains about Senator Wellstone’s recent comment to the press that he didn’t know that he was not paying into social security.

1 D: Wellstone went into a big tirade
 2 that he was not a liar
 3 and then
 4 a little while later
 5 somebody asked him about
 6 why aren’t the federal employees paying into social-
 security
 7 B: <<all> he [said he didn’t know]
 8 D: [and he said
 9 I don’t know they weren’t
 10 (.)
 11 B: well you don’t believe him
 12 D: no I don’t believe him
 13 either he was not telling the truth
 14 or he is <<force> stupid>
 15 not to know
 16 that he is not paying into social security as a [senator
 17 B: [well-
 you can call the man <<force> stupid>
 18 if you want to call the man <force> stupid
 19 now he was not <<force> tak>ing
 20 any heat from anybody
 21 you know somebody called him something else
 22 and he really went after him

23 the fact is
24 → <<stylized> maybe the guy didn't know>
25 P: maybe he's naive

In this excerpt, both Carlson and the caller, David, speak with various degrees of prosodic stylization. In his first turn, which is not prosodically stylized, David informs Carlson of Wellstone's misdoings. Rather than making a direct assessment of Wellstone's activity he simply reports what happened. David carefully constructs his turn with the logical implication that Wellstone is a liar, but he does not deliver the actual assessment. It can be argued that David is constructing his turn so that the negative assessment of Wellstone does not come from him, but rather from Carlson. In line 10 is a longish pause where such an assessment would come. Carlson, however, rather than delivering the negative assessment or showing any kind of agreement, jumps the beat established by David and quickly comes in with the information that David was building up to: "he said he didn't know" (line 7). This occurs in overlap with the slow delivery of David's utterance, which he gives in direct speech. Carlson thereby "steals David's fire": the climax which he is building up to loses its impact, and instead of giving the negative assessment of Wellstone, which would show agreement with David, she simply says what David says but in a much more deflated way than David's utterance.

Carlson's disagreement becomes even more apparent in her subsequent turns. Following her question in line 11: "well you don't believe him", David again attempts to make his point that Wellstone is a liar without directly making the assessment. In lines 13-15 he offers an argumentation figure known as "bifurcation" or "Dilemma/Fallunterscheidung"³⁰: "either he was not telling the truth or he is stupid not to know that he is not paying into social security as a senator." Except for prosodic highlighting on the syllables "stupid", which are realized with extra force and with a musical interval, David's turn is not stylized. As in his turn above, David's turn here strongly implies that Wellstone is a liar, but he does not call directly call him a liar. The logical implication is, of course, that Wellstone is a liar, and as above, it can be argued that he is constructing his turn so that Carlson makes the natural logical

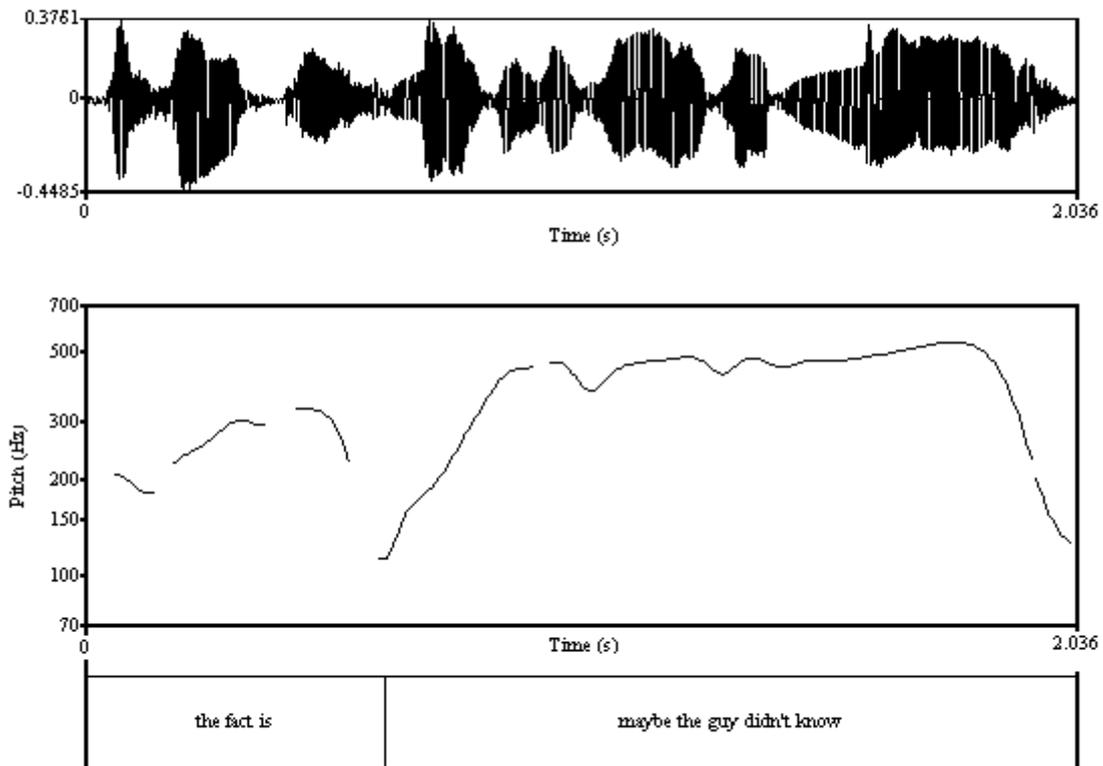
³⁰ In this argumentation figure, a speaker proposes two possible choices, one of which must be true: if the one choice is obviously not true, then the other choice must be true. See Schleichert, 1999:29-30, 177.

assessment which follows from his informing. Rather than giving an assessment in line 17, however, Carlson, in an uncooperative way as in line 7, interrupts David's turn. The beginning of this turn is prosodically stylized: "well you can call the man stupid if you want to call the man stupid". Carlson's "stupid" in lines 17 and 18 are highlighted similarly to David's "stupid" in line 14: the syllables have extra force and have a musical interval between them. She is orienting not only towards the word, but also to its prosodic highlighting. This ornamenting of the syllables with extra force is heard again on the syllable "**taking**" in line 19. Carlson not only duplicates David's prosodic stylization, she goes beyond it. The syntactic parallelism and isochronous rhythm in lines 17 and 18 make Carlson's contribution more stylized than David's utterance before. This contextualizes her dissent with David's implied assessment that Wellstone is a liar.

In lines 19-23, except for the highlighting on "**taking**" in line 19, Carlson speaks in a non-stylized way. Indeed her informing sounds as if she is taking over David's way of logically arguing: "now he was not taking any heat from anybody somebody called him something else and he really went after them". The logical way of arguing ends in line 23 with her utterance: "the fact is". The expectation that a logical fact is going to be uttered is broken in line 24, which is extremely stylized in the manner described above, and – compared to the relatively low pitch range of the previous utterances – is realized extremely high in Carlson's pitch range: approx. 500 Hertz.

The following frequency analysis underscores the perception of stylized pitch:

3.3.1.6



The break in expectation of the logical argument which would follow Carlson's line "the fact is" is reinforced by the break in prosody as well as the break in formality. Carlson refers to the Senator as "the guy" and no longer as "Senator Wellstone". With the help of prosodic stylization Carlson is able to express that although it may be stupid of Wellstone not to know that he is not paying into social security, it does not really matter: Wellstone is not a liar and stupidity is only human. Wellstone is not to be damned for his ignorance.

Following Carlson's stylized utterance in line 23 co-moderator Peter Theo says the same thing as Carlson, and also in a prosodically stylized way: "maybe he's naïve". This is heard as an orientation to Carlson's prosody in the preceding line. In both cases Carlson's assertion that Wellstone "didn't know" and Peter's assertion that he

is “naïve” are heard as downgrades from the much stronger assertion by David that Wellstone is a “liar.”

The above example shows that prosodic stylization is interactionally relevant to the participants themselves, who also display an orientation to the pattern.

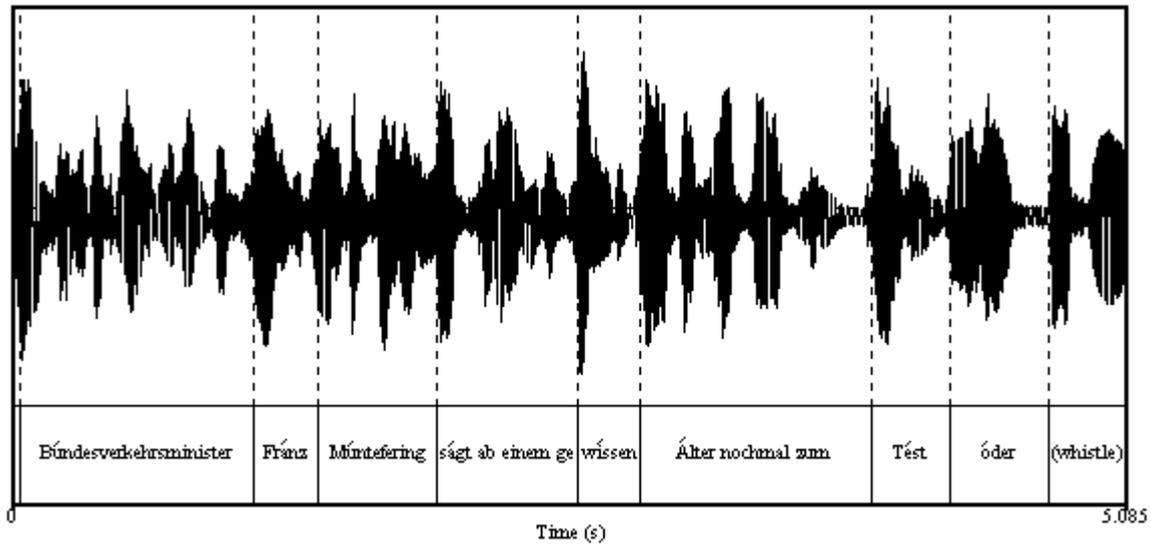
3.3.2 Gestalt Reduction in the Rhythm of Prosodic Gestalts

In stylized speech not only can the pitch be reduced within the original gestalt – as demonstrated above - so can the rhythm. English and German are stress-timed languages, i.e. in contrast to syllable-timed languages such as French, the stressed syllables occur more or less at equal distance from each other. In unstylized speech, however, there are many irregular local variations in timing which prevent the perception of an underlying rhythmic gestalt of the speech utterance. In stylized speech, however, these variations may be suppressed, resulting in an isochronous rhythmic gestalt that is, auditorily, highly stylized.

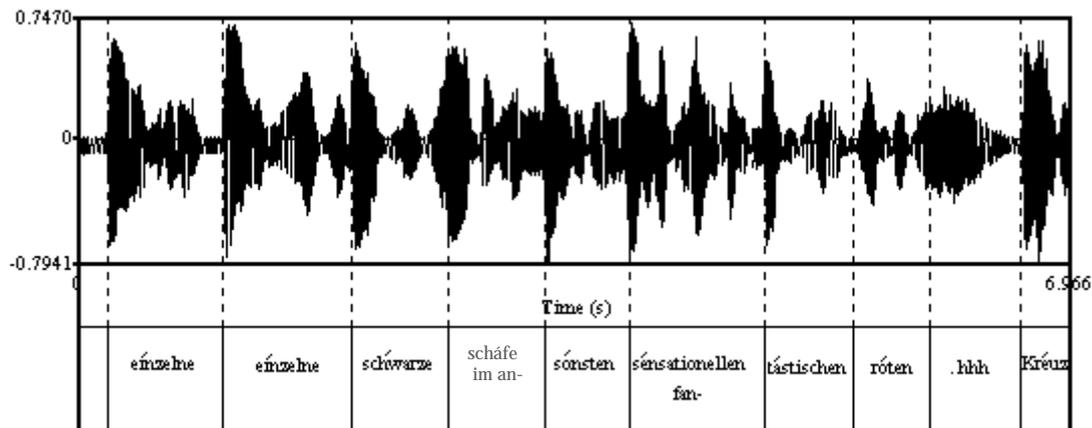
According to Couper-Kuhlen (1990), an isochronous rhythmic gestalt becomes established when at least three prosodically prominent syllables are perceived as occurring at regular intervals in time. Any number of unstressed (or secondarily stressed) syllables may occur between the stressed syllables as long as the perception of the isochronous rhythm remains intact. The stressed syllables then become like the beat in music, and the time between stressed syllables like the measure. In an isochronous rhythmic gestalt one can easily tap out the resulting rhythmic gestalt.

An acoustic analysis can help support the impression which one perceives auditorily. Consider the following two wave forms of unstylized and stylized rhythmic gestalts. Both excerpts are taken from the same monolog from the German late-night TV comedy host, Harald Schmidt:

3.3.1.6 Unstylized Rhythmic Gestalt: “Nochmals zum Test”



3.3.1.7 Stylized Rhythmic Gestalt: “Einzelne schwarze Schafe”



Compared to the wave form of the unstylized speech segment, one can see how the rhythm of the stylized speech segment is isochronously organized into a stylized gestalt. The time between the stressed syllables, which is irregular in the first example 3.3.1.6, is regular in the second example 3.3.1.7. This can be seen in the wave forms above, in which the beginnings of the stressed syllables are marked by vertical dotted lines. In the unstylized wave form, the distance between the beginning of the stressed syllables is 1.1 seconds, 0.3 seconds, 0.5 seconds, 0.7 seconds, 0.3 seconds, 1.2 seconds, 0.3 seconds, 0.4 seconds. In the stylized example the

distance between stressed syllables is 0.8 seconds, 0.9 seconds, 0.7 seconds, 0.7 seconds, 0.6 seconds, 0.9 seconds, 0.7 seconds, and 0.8 seconds.

Local variations, which prevent the perception of a stylized gestalt in the first example, are regularized, and the result is a stretch of speech which is highly stylized. Clearly, the strong rhythm, as well as other factors, such as alliteration, is what prompts Schmidt to make the meta-comment in line 19: "Das habe ich schön gesagt" ("*I said that beautifully*"). In the acoustic analysis of the stylized gestalt, the broken vertical lines at the beginning of the stressed syllables fall at isochronous moments in time, approx. 0.7 seconds from one another. In the first example, the stressed syllables (which are more difficult to discover) are not isochronous; moreover, there is no perception of a rhythmic gestalt.

The above section, 3, asserts that reduction is an important stylization process. Reduction, however, is not the only process of stylization. In the following section, 4, I assert that ornamentation is also a stylization process that is used in order to reinforce certain elements of the underlying gestalt.

4. Stylization in Visual, Musical, Prosodic Gestalts: Gestalt Ornamentation

The stylization of a visual, musical, and prosodic gestalt through the process of ornamentation can also occur on the global level. The following section will show how more complex gestalts may be stylized through ornamentation. Stylization through global ornamentation may occur through two processes: 1) an additional gestalt feature may be laid over a larger gestalt; 2) repeated local variations may bring out a larger gestalt pattern. As demonstrated above, for the ornamentation of a gestalt local irregular variations may first be reduced before the macro-gestalt is stylized through ornamentation. As in the above sections, I will demonstrate the stylization of macro-gestalts through ornamentation by offering first visual and then musical examples. Finally, I will show that such ornamentation is equally applicable to prosodic gestalts.

4.1 Global Gestalt Ornamentation in Visual Gestalts

Ornamentation of visual gestalts at the global level occurs when the global gestalt is marked either through the overlay of an additional gestalt parameter or through the overlay of patterned local variations which are spread across an entire gestalt structure. For a demonstration of stylization in which a new gestalt parameter is overlaid on top of the gestalt, consider the following visual gestalts, 4.1.1 and 4.1.2:

4.1.1 Reduced Visual Gestalt:



4.1.2 Globally ornamented visual gestalt through overlay:



Gestalt 4.1.1 shows a reduced visual gestalt, in which irregular local variations have been suppressed. Gestalt 4.1.2 demonstrates the further stylization of the gestalt through the overlay of an additional gestalt parameter.

Not only can a gestalt be stylized through the overlay of an additional gestalt parameter, but it can also be stylized when a local variation is repeated over the entire gestalt structure, as demonstrated in the following visual gestalt 4.1.3:

4.1.3 Ornamented Global Gestalt: Patterned, Local Variation



That a gestalt can be ornamented on the global level through the overlaying of an additional gestalt parameter or through the repetition of a local gestalt, which becomes a pattern, can also be seen in the following musical gestalts.

4.2 Global Gestalt Ornamentation in Musical Gestalts

The two processes for ornamenting a global gestalt through the addition of a single gestalt parameter or through the repetition of a local variant across a global gestalt structure can also be applied to auditory temporal gestalts. The following will show these two processes applied to a musical context. For this purpose, consider, once again, the Irish jig, “Tripping Up The Stairs”:



Through the overlaying of an additional gestalt parameter, a musician can draw attention to a certain segment of a larger gestalt. Volume is often used for this purpose, as one can add it over the entire gestalt structure. With volume, by playing

louder or softer, a musician is able to mark a larger gestalt structure, such as in the following example:



By changing the volume of the second measure, a musician can draw attention to this part of the gestalt structure. Although this gestalt is part of a larger gestalt, one can speak of global ornamentation as smaller gestalts are nested within the more global structure. The grouping of the gestalt can be seen as following:

(₁(₂XXXXXX₂)(₃XXXXXX₃)(₄XXXXXXXXXXXX₄)₁)

- 1 entire gestalt
- 2 volume (forte)
- 3 volume (piano)
- 4 volume (forte)

A larger gestalt structure be ornamented not only through the overlay of volume, but through para-musical elements as well, such as timbre. A fiddle player, for example, might choose to draw attention to this same gestalt by bowing closer to the bridge, making the sound of the violin harsher and scratchier.

Collaboratively, two musicians might draw attention to the same gestalt when one musician plays the entire section and the other plays only the second measure, thereby overlaying her contribution to the gestalt structure.

Whereas volume and timbre are laid over a global gestalt structure in order to ornament it, a musician might also choose to add patterned local variations across a gestalt structure. Consider the following example:



The first note of the gestalt has increased volume. This local gestalt manipulation is repeated for the fourth note, the seventh note, and the tenth note. The repetition of the local gestalt parameter groups the first two measures together in an ornamented gestalt. The manipulation is not applied to every single element within the gestalt, but only every third element. The grouping can therefore be depicted as follows:

(1(2(3X3)XX (4X4)XX (5X5)XX (6X6)XX2) (7XXX XXX XXX XXX7)1)

1 entire gestalt

2 stylized gestalt through patterned local variation

3,4,5,6 forte

7 unstylized gestalt

A local gestalt structure which is manipulated in a certain way, and which occurs in a pattern with other similarly manipulated local gestalts is able to ornament the global gestalt structure in which the manipulated local gestalts occur. Such manipulation may be through volume, pitch, duration, or any para- or nonmusical gestalt features.

The following will now demonstrate the stylization via ornamentation of global gestalts in prosodic structures.

4.3 Global Gestalt Ornamentation in Prosodic Gestalts

Just as in visual and musical gestalts, prosodic gestalts may be stylized through ornamentation on the global level through overlaying and through patterned repetition of local variations. The following section will show the stylization process of ornamentation through these two strategies.

4.3.1 Ornamentation through Patterned Repetition of Local Variations

Similar to a musician playing with musical gestalts, a speaker in conversation may choose to draw attention to a certain segment of speech by repeating a local

variation over a longer gestalt structure. The effect of such repetition of a local variation is a stylization and a grouping of the elements which fall within the repetition. Prosodic as well as non-linguistic gestalt features laid over local gestalts and then repeated over a more global gestalt ornaments the global gestalt. As in music, one non-linguistic way that a more global gestalt may be brought out is through a percussive tapping, slapping, snapping, or clapping of hands, pencils, etc. As in music, a global gestalt can be created through such percussive accompaniment.

In the following example, the speaker B overlays a percussive tapping over local gestalts. These are marked in bold and set off by parentheses in the transcript. The repetition of this local gestalt ornamentation stylizes this segment of speech and groups the various gestalts into a unified global gestalt:

4.3.2 “Dead Boyfriend”

Weather broadcaster, Valerie (V), chats with Barbara Carlson (B) before reading the weather forecast. The theme of Carlson’s show has been death and dying, and Valerie has just summarized a book written by a woman who claims to have died and returned to life. Barbara explains how she talks to deceased friends and relatives.

1 V: no I mean its backed up by (.)
 2 several accounts of people who have
 3 [died and (.)]come back
 4 B: [who died [yeah I believe you can do that
 5 and I be(**lieve**) there are people over there
 6 and (**I**)
 7 let me tell you
 8 I(**talk**)all the time
 9 .hh I have an old boyfriend that (**died**)
 10 uhh about six months (**ago**)
 11 and I can’t tell you the conversations I’ve had-
 with him
 12 V: oohh
 13 B: ab(**so**)lutely incredible

When local irregular variations in rhythm are first reduced, resulting in an isochronous rhythmic pattern, and then overlaid by such percussive slapping, tapping, and clapping one perceives a stylized gestalt. Consider the following example, in which Bill Clinton, during the 1996 presidential debate with Republican candidate Bob Dole,

stylizes a segment of speech first through the reduction of irregular rhythmic variations, and then through ornamentation via the overlaying of clapping on local nested gestalts occurring at isochronous intervals. As above, claps are marked in the transcript in bold print:

4.3.3 “Issue by Issue by Issue”

In the 1996 presidential debate against Republican candidate Bob Dole, Bill Clinton (B) discusses US financial problems with China.

1 B: there is no easy way to do this
2 when you're dealing with economy that's traditionally-
been more close
3 and one that's been traditionally been more open
4 you just have to gut it out (**iss**)ue by (**iss**)ue by (**iss**)ue

Not only can the repeated percussive tapping or clapping be added to local gestalts for the ornamentation of more global gestalts, other gestalt features may similarly be added to local nested gestalts and then repeated for the same effect of global ornamentation.

5 Summary

The preceding chapter shows how, from a gestalt perspective, prosodic structures, like musical and visual gestalts, may become stylized. “Highlighted” and “stylized” gestalt patterns are proposed to describe how speakers may highlight a single part of a gestalt structure (highlighting), or an entire gestalt structure (stylization). Three gestalt processes are proposed for describing how speakers highlight and stylize their spoken contributions: 1) reduction, 2) adjustment, 3) ornamentation. Reduction is the suppression of local variations, resulting in what von Ehrenfels refers to as a “pure Gestalt”. Adjustment is the regularization of irregular gestalt features into patterned ones. Ornamentation is the overlaying of additional gestalt parameters on top of an underlying gestalt.

This chapter concludes the formal/descriptive section of the dissertation. In the following I shall turn to questions of function. In this section prosodic stylization is seen to be a discourse practice which interactants use for specific goals within interaction.

Chapter 4: The Placement and Stylization of *oh* in Informing Sequences

Whereas the first part of this dissertation describes from a gestalt theoretic perspective the perception of grouping patterns and stylized prosodic gestalts in spoken discourse, the following chapters will now take an interactive, decidedly CA approach to show how prosodic stylization and the specific placement of a gestalt within another gestalt is interactionally relevant. I will show that these two phenomena are discourse practices which are exploited within informing sequences to control turn taking and floor holding rights.

The act of informing has been described by Heritage as follows: “tellers propose to be knowledgeable about some matter concerning which, they also propose, recipients are ignorant” (1984:304). Following such an informing act, recipients often respond by producing the particle “oh”. In Heritage’s words: “with the production of “oh”, recipients can confirm that, although they were previously uninformed on the matter at hand, they are now informed” (1984:304). In the following I will argue that the prosodic stylization and the placement of the “oh” as a nested or complex gestalt has a specific influence on taking turns in the unfolding interaction. I will describe how interactants stylize “oh” and then place it within the gestalt of the first pair part of an informing sequence in order to achieve specific interactional effects. Moreover, I will explain why these two practices are particularly suited for accomplishing the goal of managing turn taking and floor holding.

1 Introduction

In much of CA work, a finding often begins with a so-called “noticing” after listening to long hours of data (Schegloff 1996). For the following it began with a simple observation as I was reading John Heritage’s 1984 article “A change of state token and aspects of its sequential placement”. I noticed that many of the *ohs* which occurred in the transcriptions which Heritage used were lengthened and that they

often occurred in overlap with the first turn informing³¹. Although these two phenomena are transcribed in Heritage's data, he does not comment on them in the article. Instead, he focuses his attention on *oh* as a "state change particle".

John Heritage argues that in spoken discourse, participants express a change of "locally current state of knowledge, information, orientation or awareness" (1984:229) through the particle *oh*. He demonstrates this claim by analyzing *oh* in two kinds of environments: informing sequences and repair sequences.

In informing sequences *oh* – in contrast to "yes" or "mm hm" – is used to show that the information which was offered in the preceding turns has been informative to the listener. According to Heritage, "*oh* is used to mark the receipt of the informing delivered in the preceding turn or turns. Moreover we can additionally notice that these *oh* receipts (1) occur in response to complete chunks of information and (2) are produced at points at which the informings are possibly complete" (301). The following examples, taken from Heritage's article, demonstrate these claims:

"Things 'av arrived"

I: Ye:h. .h uh:m (0:2) I've jis' rung tih teh- eh tell
you (0.3) uh the things 'av arrived from Barker'n
Ston' ou[:se,
→ J: [Oh::::.
(.)
J: Oh c'n I c'm rou:nd, hh

"Dez's Mum's "

J: =Hello there I rang y'earlier b'tchu w'r ou:t,
→ I: ↑Oh: I musta been at Dez's mu:m's=
→ J: ↓=aOh::. h=

(1984:301)

³¹ Heritage's data (as well as my own in this chapter) were transcribed by Gail Jefferson and retain, for the most part, her original transcription conventions.

In these examples the particle *oh* immediately follows an informing. The *oh* can be free standing, such as in example one and in line three in example two, or it can precede additional turn components, such as in line two of the second example, where the *oh* is followed by a justification for why the speaker was out. When a turn containing *oh* is not freestanding, then it may, as Heritage asserts, not only mark a change of state in a person, but, additionally, show the speaker dealing with the information. Often speakers are required to deal in some way with the information that is given. Speakers may, for example, include in their turn an assessment following *oh*, as in the following:

“Takin’ a Train”

A: ... Well lately in the morning Rosemary’s been
picking me up. -Yihknow so I (haven’ been) even
takin’ a train in [(the morning)
→ B: [hhOh that’s grea:t.

(1984:302)

Another activity following the change of state token might be to request more information, such as in the following:

JG:3C:5

R: I fergot t’tell y’the two best things that
happen’ tuh me t’day
→ C: Oh super.=What were they

(1984:303)

In the above examples the particle *oh* shows that the information in the preceding turn is received as new by the second speaker. Furthermore, within the turn which contains the particle, additional turn components can be added to assess the information or request additional information, for example.

One aspect which Heritage does not consider but which is taken up by John Local (1996) is the role of conversational phonetics in informing sequences. Local attempts to show “the extent to which phonetic parameters are intertwined with lexis and

syntax in the interactional functioning of *oh*" (1996:178). Local shows that such phonetic parameters are oriented to by the interactants and are exploited for achieving conversational effects. Some specific claims by Local will be described as the chapter progresses.

My observation, that many of the *ohs* are stylized and systematically placed within the first pair part to accomplish certain interactional goals, complements both Local's observation that conversational phonetics influences interaction as well as Heritage's observation that *oh* is used by interactants as a particle which signals a "state change" in the speaker after hearing new information. Following my initial observation, I had the following questions:

- 1) Why, and in what contexts do speakers stylize their *ohs* in informing sequences?
- 2) When do they place their contributions in overlap? Is there an interactional relevance for placing the *oh* within the first turn of an informing sequence?
- 3) Are these two phenomena "practices" in Schegloff's sense? And if so, what "actions" are they accomplishing?
- 4) Why do interactants explicitly exploit these particular practices in order to achieve these results following an informing?

To answer these questions I put together a collection of examples of the suspected phenomena from the Newport Beach corpus, which consists of 21 taped telephone conversations. After closely analyzing the data, it appears that within informing sequences interactants do exploit prosodic stylization and the specific placement of *oh* in relation to the first turn as actions in order to control turn-taking. How and why interactants do this will be described in the following.

I will show that interactants who are news recipients face a delicate situation. They must, on the one hand, show that the news has caused a so-called "state change". On the other hand, however, they must also not interrupt the first speaker or, worse, take over the floor during a news informing. It will be shown that stylization and placement help accomplish this task. It will be shown that various kinds of stylization signal to the first speaker that s/he should continue. Non-stylization shows the

second speaker taking over the floor and continuing the next action. Similarly, a nested *oh* which occurs in overlap with the first turn informing early in the sequence also signals to the first speaker that s/he may continue unhindered. A complex *oh* which begins before the first turn is over and continues till after the first turn is completed signals that the second speaker is going to continue with the next discourse action.

The following is organized into three sections. I will first show *ohs* which are not in overlap but which directly follow the first turn informing. The specific prosodic stylization on these freestanding *ohs* signals to the first speaker that s/he may continue the next action. I will then show *ohs* which are stylized and which occur as a nested gestalt early within the first turn informing. Again, the specific kind of prosodic stylization coupled with the placement of the *oh* signals to the first speaker that she should continue. Finally I will present *ohs* which start before the first turn ends and continue beyond the first turn in what I call a complex gestalt. These *ohs* signal that the second speaker should continue the next action.

2 Non-overlapped oh

The first case I will describe shows *oh* directly following an informing without overlap. John Heritage (1984: 302) and John Local (1996:180) both assert that a free standing *oh* following an informing is rare, although in my data I had little difficulty assembling a collection of examples. This first section is divided into three parts, all of which show *oh* directly following the first turn informing without overlap. In the first part I will show free standing *ohs* and describe how interactants orient towards their placement within the gestalt of the first turn informing and to their prosodic stylization and the effect this has on floor holding and turn taking in the unfolding interaction. In the second part, I will show *oh* followed by additional turn components, such as an assessment, for example. It will be shown that speakers stylize these contributions in a similar way as with free standing *ohs* with an additional harmonic interval between the *oh* and the assessment. Finally, in the third part, I will then present “deviant” cases in which second speakers continue with the next activity after saying *oh*. In

such cases, speakers do not stylize their contributions as much as with free standing *ohs* and *ohs* followed by an assessment.

2.1 Freestanding *oh*

The data systematically show that freestanding *ohs* which are not in overlap have a specific prosodic make up: they are lengthened and lower in volume than the previous talk and the following talk. The sequential development of such free standing *ohs* is as follows:

Speaker 1: informing
(optional pause)
Speaker 2: *oh*
(optional pause)
Speaker 1: new activity (or, as is often the case, another related informing)

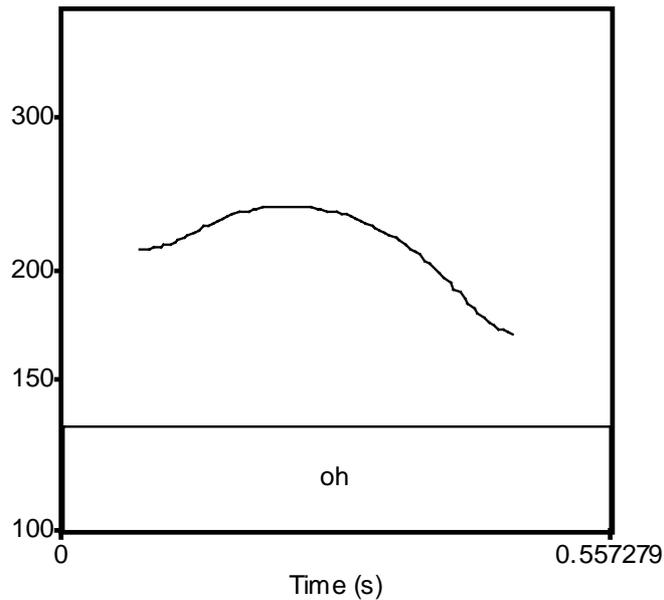
This may be seen prototypically in the following examples:

NP1, Tape 8, 14.13

1 → Nan: so: I jist assoom that the notice the e: the=
=telegram that went fr'm th'bank w'ss return' becuz he
didn't w:ant to accept it.
(0.4)
5 → Emm: OH: .h
→ Nan: °See, °
Emm: °Mm hm? °

In lines 1-3 Nan informs Emm about her current situation with her ex-husband. This is followed by a pause in line four. Consistent with Heritage, Emm follows Nan's informing with an *oh* which has a falling pitch contour. This is followed in line 6 with Nan's next action, "see".

The contour pitch contour of Emm's *oh* can be seen in the following frequency analysis:



The analysis shows a specific falling pitch contour that systematically occurs in all the examples of the collection. Following a first turn informing the second interactant produces this pitch contour with a free standing *oh*. The *oh* is then followed by a new action by the first speaker.

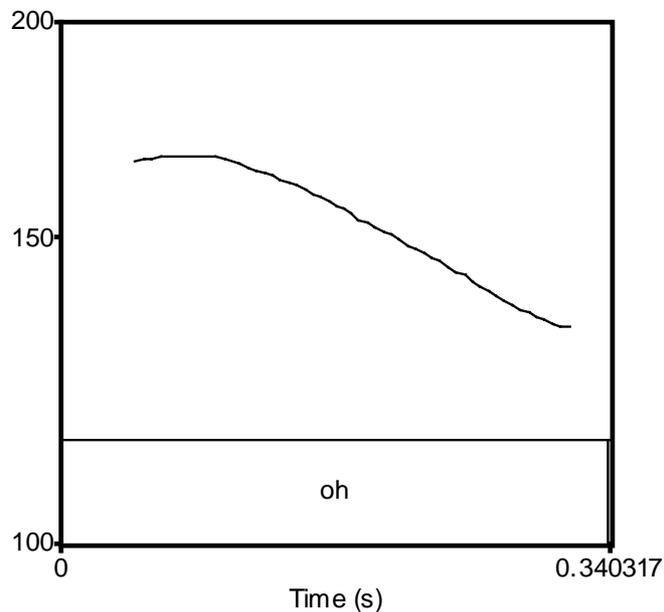
Often the new action by the first speaker is another informing. Speakers can create an “informing chain” by giving an informing, followed by the particle *oh*, which is followed by another informing, which again be followed by the particle *oh*, etc. Such an informing chain can be seen in the following example:

NP1, Tape 6, 1.09

- 1 Lot: .hhh Hey I gotta good trip if he wantstuh go it's
 onna Mondee though c'n 'e go on tha[:t?]
 Emm: [n:] No=
 → Emm: =u- he's goin with the Gas Comp'n eez all booked
 5 up. So that'll be enough fer him.
 (0.5)
 → Lot: Oh:.
 → Emm: Ther goin ont yihknow they (.) charter a boat f'm
 San Diego.

In lines 3-4 Emm starts with an informing. This is followed by a pause and then Lot's “oh”, which is followed by another informing in what I have termed an “informing

chain". A frequency analysis shows a similar contour as in the first example (NP1, Tape 8, 4.18) above:



Following is another example of an “informing chain”:

NP1, Tape 9, 1.49

- 1 Lot: We:ll? euh: (0.3) in the morning we fi:shed ed the: uw- uh
fished on the ti::de (.) uh high tide we fish' ba::ss in
 the bay th-there jis et the mouth a'th'jetty.
 Emm: Mm:: hm[:
- 5 Lot: [.hh Then: (.) we went down tuh: whur we caught the
big ha:libut yihkno:[w e-]e- u]h en:: uh .hh we wo:rked=
 Emm: [Mm:] h m]
- Lot: =THA:T there'n we fi:n'lly go:t'm up real close tih the
ro:cks.
- 10 Emm: Oh:::
 → Lot: Bah:t (.) five'n six pou:nders.
 Emm: Oh: goo :::[d.
 → Lot: [Bud you gotta work tha:t spo:t in there beef
 uh yihknow: tuh know exa:c'ly whur they a:re.=

Following Lot's informing in lines 8-9 is Emm's *oh*, which is followed by another informing, which is followed by an *oh*, which is again followed by another informing in an informing chain.

Local observes that a topic change frequently follows such an *oh* (1996, 182). Again, these changes of topic often take the form of an informing. Consider the following, for example:

NP1, Tape 7, 0.46

- 1 Emm: =[Bud's gon'play go:lf now up Riverside he's js leavin'
(0.2)
Lot: Oh:.
(0.5)
- 5 Emm: So: Kathern' Harry were spoze tuh come down las'night
but there wz a death'n the fam'ly

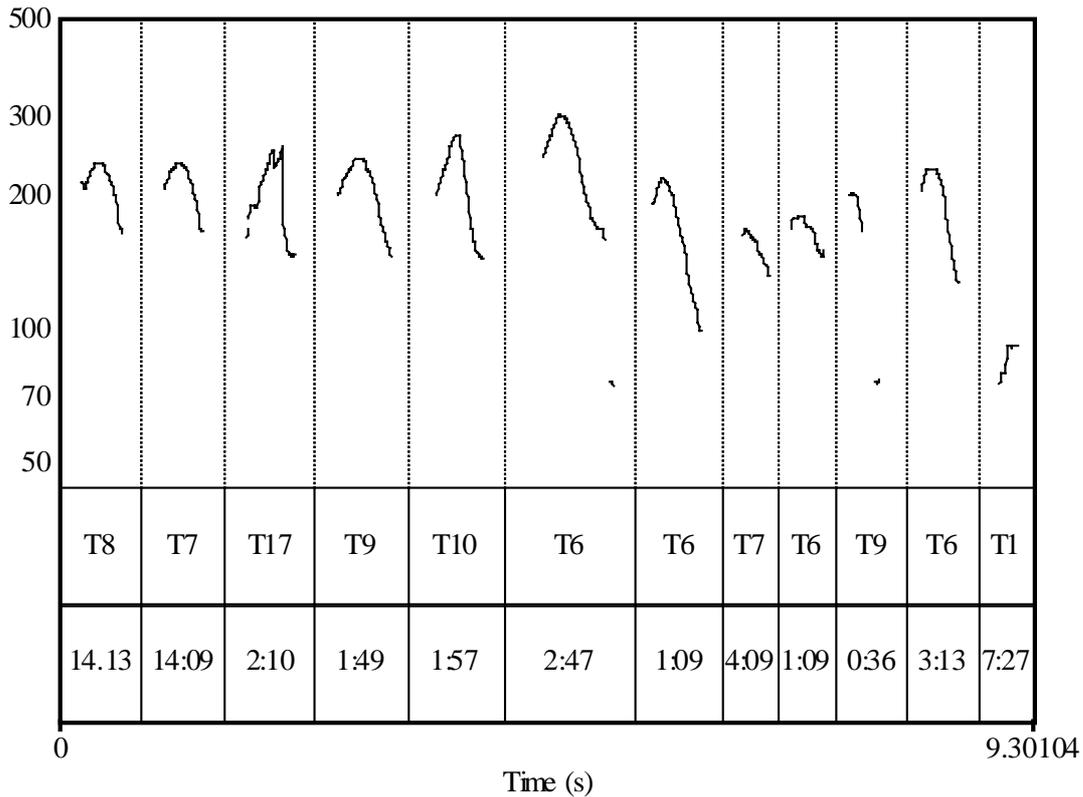
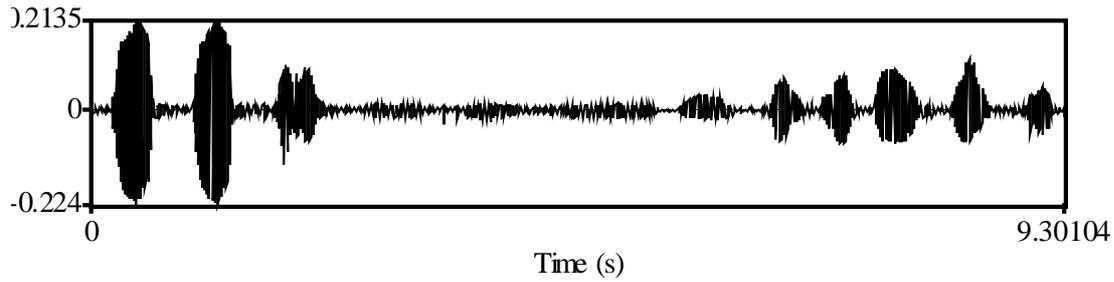
In line 1 Emm informs Lot about Bud playing golf. This is followed by a pause and a free standing *oh* with the specific pitch contour described above. Following another pause, Emm begins a new informing, interestingly beginning the new activity with the discourse marker “so”. This can be seen as an indication that free standing *ohs* are used to give the floor back to the first speaker, who is then to continue in some way, either with a further informing of the topic in hand, or with a new topic.

Another example of syntactically continuing the second informing with a discourse marker can be observed in the following, this time with the conjunction “but” rather than “so”:

NP1, Tape 14, 0.38

- 1 Lot: How's Mister Bla:ck.
Emm: We::ll Barbra siz'e had the look a'death onnim ah don'know
he's gotta (.) pro:strate (0.2) .hhhh glan'going uh (.)
backing up in iz khidney: hh
(0.2)
- 5 Emm: .hh He's ho::me:
(0.2)
Lot: Oh::.
(0.2)
- Emm: But uh: she sid God'e look' •t*errible hhhh
Lot: Ye::uh,

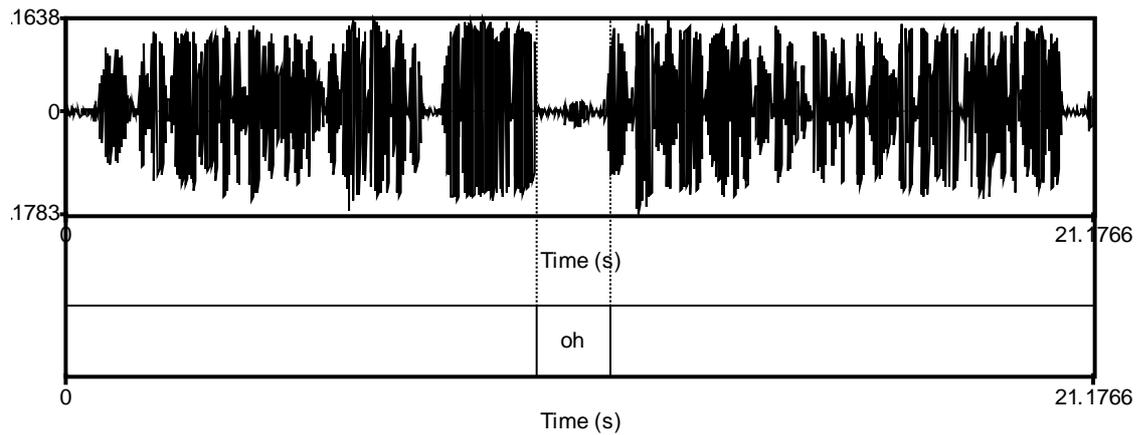
In all of these examples of free standing *ohs* not in overlap there is a specific prosodic contour which can be seen in the frequency analysis of the following examples.



Not only do the *ohs* show this specific pitch pattern, they are also often reduced in volume. Consider for example, the following:

NP 1 , Tape 10, 1.57

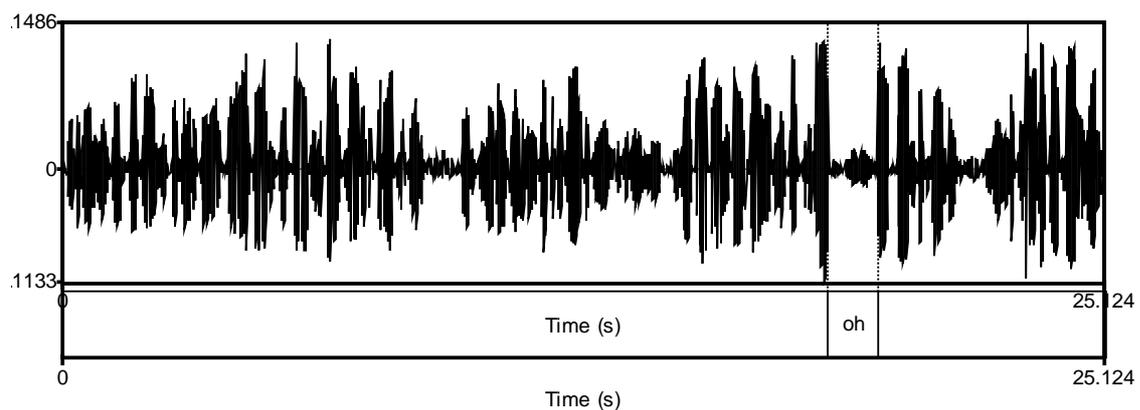
- 1 Emm: .t.hh I've dever had a skin blemish in mah •li:fe yihknow
 en I got this stuff ennit eeyuh ih •you•:'ve ih You •SEE IT
 ADVERTI:ZE'the PA•P*ER.
 (0.3)
- 5 Emm: P-S-O-R-I-A-S-I-S, .hhhh[h h h h] .hhh
 → Nan: [.t °Oh:] : .°
 Emm: An' i-ih sca:les 'n comes o:ff thez: in a lid'le bid of
 i:tchi:ng? So'e gay me s'm saav 'n then'e looked it thet
 toe'n'e siz God that's gotta come ah:ff becuz
 10 [I've been maa- messin a]round w*ith *it.



The lower volume can also be seen in the following analysis:

NP1, Tape 9, 0.36

- 1 Emm: Getcher fam'ly the:re?
(0.5)
Lot: Uh-no:: Sa:m le:f' °cut that fish up (.) fer th-the ca:t
th'rest of it,° is
5 (0.4)
Emm: Did'ee?
(0.3)
Lot: Eh yeah 'e left Mondee.
→ Emm: Oh:.
10 Lot: Yeahm.
(0.2)
Emm: Mm hm O:kay honey well I know yer busy [I jst]



The question remains: why do interactants exploit these two practices for this specific social action? Why does the recipient of an informing prosodically form the free

standing *oh* with this pitch contour and with a lower volume than the surrounding discourse?

I claim that the pitch contour with its steeply falling pattern and the low volume iconically signals “I am not going to continue, you continue”, and, indeed the first speaker does continue in all the cases I could find. John Local asserts that “one account for the pitch choice might be that a falling pitch contour here strongly projects finality/completeness” (1996: 183). Although the claim that a falling pitch contour is always associated with finality is refuted by Local himself (see Local 1986, Local, Kelly and Wells 1986), the claim does seem to have a certain validity in informing sequences.

Within informing sequences falling pitch is associated with finality, lower volume indicates non-competitiveness. The second speaker is making no claims to the floor. The action that these two practices help accomplish is one of turn-taking. The *oh* indicates state change but immediately and unequivocally gives the floor back to the first speaker, who, more often than not, continues the next action with a new informing.

It seems that controlling the floor in this way is closely associated with the very nature of informing sequences themselves. A recipient of news often feels compelled to offer a back channel, and, indeed to show a change of state, on the one hand, but she might not want to cut off the informing, particularly if it is a so-called “hot news informing”. The way this is accomplished is to quickly offer the *oh*, but in such a way that the first speaker immediately continues the informing. For this interactional goal, reduced volume and a final pitch contour seem optimal practices.

2.2 Oh followed by additional turn components

Heritage and Local both show that an *oh* following an informing may contain additional turn components, usually an assessment, but sometimes a “next-utterance-soliciting component” (Local 1996: 186). Consider the following examples:

NP1, Tape 8, 0.23

1 Emm: .hh How you doin.
 Nan: .t hhh Pretty good I gutta rai:se.h .hh[hh
 Emm: [Kuu:u[d.
 Nan: [Yeh two dollars

5 a week..h
 (.)

→ Emm: oh [w o : w .]
 Nan: [uh::: h]uh hu[:h hu:h]
 Emm: [Wudee gun:] do with it a:ll.

10 Nan: Gol' I rilly I jis don't know how ah'm gunnuh spend all that
 money.

Following an informing in lines 4 and 5 and a pause in line 6, Emm says “oh” and follows it with an assessment “wow”.

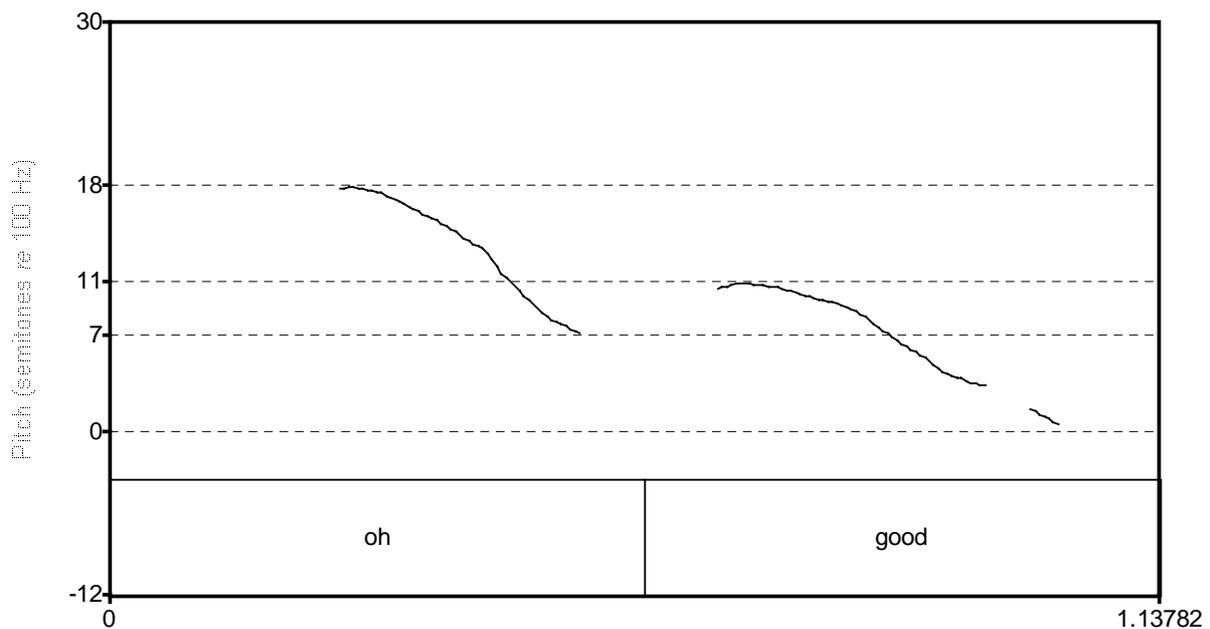
NP1. Tape 7, 0.50

1 Emm: So: Kathern' Harry were spoze tuh come down las'night
 bpt there wz a death'n the fam'ly so they couldn'come
 so Bud's as'd Bill tuh play with the comp'ny deal so I
 guess he c'n play with im so

→ Lot: Oh:: goo::d.
 Emm: WHAT A MISERBLE WEEKE:ND

Following an extended informing in lines 1-4 Lot produces the particle “oh” and follows it with an assessment “good”.

In each of these cases the *oh* and the next turn component display a prosodic production that is similar to the free standing *ohs* described above, with one difference: the two elements “oh” and the assessment are heard as producing a musical interval. An acoustic analysis of line 5 supports this perception.



In this frequency analysis the pitch is measured in semitones. The peak of the pitch “oh” is measured at 18 semitones. The peak of the pitch “good” is measured at 11 semitones. A difference of 7 semitones, or, in other words, a harmonic fifth. Similarly, the pitch of “oh” ends at 7 semitones and the pitch of “good” ends at 0 semitones, which also gives a harmonic fifth. Both pitches fall exactly 11 semitones.

The same pitch pattern with this interval occurs systematically in the examples in which an *oh* is followed by an assessment. Like the free standing *ohs*, the *ohs* with assessments as additional turn components are heard as having a similar prosodic design: pitches that fall (the falling pitches are heard to have a musical interval). Unlike the free-standing *ohs*, however, they do not display a systematic reduction of volume. A possible explanation is that assessments must often be prosodically highlighted in order to be heard as preferred (this will be discussed in detail in the following chapter). In any case, following the *oh* with the additional turn components above, the turn always reverts back to the first speaker.

I argue that these cases involve the same type of prosodic stylization as free standing *ohs*. The musical interval that occurs in these responses comes about in conjunction with the extra syllable. The stylization here has the function of signalling to the other recipient “my turn is complete”.

2.3 Deviant Cases

I will now present some deviant cases to the assertions made above and offer some possible explanations for why I see them as deviant. The following are examples in which *oh* is followed by a next action by speaker two rather than speaker one.

Consider the following sequence with a free standing *oh*:

NP1, Tape 6, 0.52

- 1 Lot: He went out marlin fishing last night.
Emm: Oh:.
(.)
Emm: Cz ther: gittin s'm albaco:re,

In this example the *oh* occurs in line 2 following Lot's apparent informing „he went out marlin fishing last night“. The *oh* is realized similarly to the free standing *ohs* described above. Unlike the examples described above, however, Emm continues her turn in line four “cause they're getting some albacore.” At first sight this seems like a deviant case, but consider the few lines before:

NP1, Tape 6, 0.52

- 1 Emm: Ah'll be da:rn.=We'll: anyway ihz ihz not too co::ld,
Lot: oh it's w:wa:rm don'tchu think it i[:s?
Emm: [Yeah. I thought maybe
Earl wz out albacore fishin.
5 Lot: He went out marlin fishing last night.
Emm: Oh:.
(.)
Emm: Cz ther: gittin s'm albaco:re,

Seen with the surrounding context one can see that this *oh* does not follow an informing sequence, but rather a repair sequence. In lines three and four Emm says she thought Earl went out albacore fishing, but Lot repairs, saying it was not albacore but marlin fishing. Emm offers an *oh*, and then an explanation why she thought he was out albacore fishing. This example is then not a deviant case at all, but simply a social action other than an informing.

Consider the next example, which also shows the second speaker continuing the turn following an *oh* in an informing sequence.

NP1, Tape 9, 4.11

- 1 Emm: S'let's head fuh the bea:ch I didn' ev'n stay fer the
dinner.
(0.3)
- Lot: Oh::.
5 (.)
- Lot: Uh huh,
____(0.7)____
Emm: .t.hhkhhh
(0.2)
- 10 Emm: S*o: I din ev'n dre:ss •UP ah made the pretties' DRE::SS=
Lot: =[°uhh hu°]

Following an informing Lot produces an *oh* in line 4. There follows a pause in line 5 and Lot then says “uh huh” in line six. The *oh* is realized as all free standing *ohs* described above with a falling contour and lower volume. It can be argued that Lot’s *oh* did not have the desired effect of giving the turn back to Emm, so her “uh huh” attempts again to yield the floor. Following the pause in which Emm does not take over the floor as desired, another *oh*, as a signal of state change, would have seemed redundant. By saying “uh huh”, however, she is able to give up the floor, but only after a rather long pause of 0.7 seconds.

Emm’s response in line 10 is part of a typical informing chain as described above. She says “so” and gives another informing.

Two other deviant examples shed light on the practices of stylization and how they are exploited by interactants to give the floor back to the first speaker. Consider the following:

NP1, Tape 9, 2.32

- 1 Lot: En Ru:th uh: this friend a'mi:ne oh: .hhh well it (.)
e-eh sh- I let 'er stay et the. 'waiian hou:se: >over the
week<. So we're goin uh: (.) e:-eh t'morruh morning ou:t.
(.)
- Emm: Oh: good. Gunnuh rent a boa: [t? er]
Lot: [Ye:::]ah=

Emm, following an informing, says “oh” and then adds two further components: an assessment, and then a question. Unlike the free standing *ohs* and the *ohs* followed by an assessment above, this example shows speaker two, following the

assessment, taking control of the interaction and initiating the next action. In this way it seems to contradict the above description.

The fundamental difference between this case and those described above is that the *ohs* and the assessment are not nearly as stylized as in the examples above. When the recipient of an informing wants to initiate a next action themselves, they *do not* stylize their contribution. The *oh* does not fall as steeply, is not realized as a musical interval, and is not reduced in volume.

For another similar example consider the following:

NP1, Tape 7, 3.01

- 1 Emm: .h Bill wen'out onna cha:rter down tuh San Clemente
(0.5) city. proper right o:ff well they left et one uh'
clo:ck on this lank- angling club'ee got about fifty
pounds a'fi:sh.
- 5 (0.4)
Emm: .hh He more th'n that 'e siz'e gave suh much of it awa:y
but'e brought home about fifty, ba::ss'n evrything.
- Lot: (n)Oh: I know it the ba::ss'r: u.-we [.h h h
Emm: [()].
- 10 Lot: The smallest ba::ss we kep' wz one pou:n' we-we uh we:ll
(.) Mondee we we:nt uh went with this: (.) gi:rl Mary en
we wen'over on Elmer's do:ck en Faye went with us 'n=

As in the example above, Lot produces an *oh* and then continues with the next action. As above, the *oh* is not heard as having the same prosodic stylization as free standing *ohs* or *ohs* followed by an assessment: the turn displays neither the falling pitch nor the reduced volume.

In the cases in which *oh* is realized following the informing there appears to be a systematic use of prosodic stylization to indicate to the other speaker that they should continue. Often the next action initiated is another informing in an informing chain. When the second speaker does want to take over the floor and initiate the next action, their *oh* utterances do not display the same amount of prosodic stylization as free standing *ohs* or *ohs* which are followed by assessments.

In the next two sections I will turn to *ohs* which occur in overlap with the informing. In the following, such *oh* turns are formed as nested gestalts. In the final section, they are formed as complex gestalts.

3 Nested oh gestalts

Whereas the first section has shown the particle *oh* immediately following the first turn informing without overlap, the following shows the *oh* completely integrated into the first turn informing as a nested gestalt. As defined in chapter 1 (p. 54) above, nested gestalts are “those gestalts in which there are two gestalts A and B such that B occurs within the larger gestalt A”. In the following examples the second speaker overlaps her “oh” so that it is fully nested within the informing turn. I will show how prosodically fitting the *oh* into the first turn gestalt via syllable lengthening, slow pitch fall, and low volume influences the turn taking of the unfolding interaction. Unlike the examples of non-overlapped *ohs* above, it will be shown that when interactants offer their contributions in complete overlap and prosodically stylized as described above, then the floor reopens following the end of the first turn informing.

Schematically, the sequential order of such a sequence look as follows:

Speaker one: first turn informing

Speaker two: *oh* fully nested within first turn informing, i.e. starting after first turn informing begins and ending before first turn informing ends.

Next action may be initiated by speaker one or two.

In such examples the lengthened *oh* occurs in full overlap and is prototypically lower in volume and lower in pitch than the surrounding discourse. The overlapped *oh* mimics the pitch contour of the informing intonation phrase, i.e. it falls as the informing intonation phrase falls. I argue that such an *oh* is a contribution which is not heard as a turn that is making any kind of floor holding claims. It occurs simultaneously with the assessment and stops before the turn is over. In such cases, either the first or the second speaker can continue the next action after the completion of the informing.

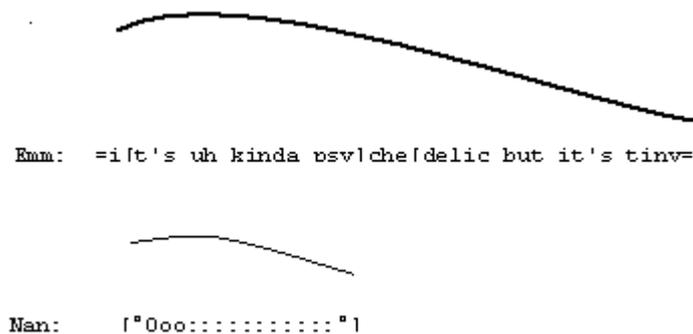
Consider the following examples:

NP1, Tape 10, 18.12

- 1 Emm: .hhh I: MA:DE ME A DAHLING DRESS tih WEAR DIH THE DESERT.God I
go[t the p-purtil'es:
Nan: [•Didju ••::h?]
(0.2)
- 5 Emm: p-print it's almos'like s:ilk but it['s euh .hhh[hh<]
Nan: [°Mm:,° [°Mm]hm,°
(.)
Emm: *u. evry color'n it's rea:l tiny=
→ Emm: =i[t's uh kinda psy]che[delic but it's tiny=
→ Nan: [°Ooo:~::~::~::~:°]
Emm: =it I mean ih-u. -ih-it u.-psychedelic isn'the word=

Nan's *oh* in line 10 is in complete overlap with Emm's informing. Following the end of the informing, Emm continues the next action.

The pitch contours of the two overlapped intonation phrases in lines 9 and 10 may be depicted as following:



Although it is lower in pitch and volume, the pitch contour of Nan's lengthened *oh* mimics the pitch Gestalt of Emm's informing in line 7 in which it is imbedded. I claim that Nan aligns the placement of her *oh* within Nan's informing for interactional purposes. By doing so, she is able to communicate her change of state "online" in the moment that the informing is being offered. By mimicking the pitch contour and placing it at a low volume and a low pitch she can make her turn coincide with Emm's without creating acoustic interference. Following the end of the informing sequence, speaker 1, Emm, continues.

In their work on assessments, Goodwin and Goodwin (1987, 1992) show that a speaker can time his/her contribution in such a way that other interactants can offer their assessments in overlap. Goodwin and Goodwin describe how interactants are able to share in the assessment by contributing their assessments simultaneously, which increases involvement between the interactants.

Unlike in assessment sequences, in informing sequences the first and second speakers do not “share” the overlapped action. In an informing sequence, only the second speaker can, of course, show a change of state to the new information. But the overlapped *oh* in an informing sequence, similar to an overlapped assessment in an assessment sequence, can contextualize a heightened involvement between the interactants. The placement of the *oh* within the first turn informing and its lengthening over a stretch of speak of the informing shows the second speaker *immediately* changing her state as the informing unfolds. There is no delay, as when the *oh* occurs after the informing is complete. It therefore shows the unmediated influence of the news on the recipient, and, in this way, also increases and contextualizes a heightened involvement in the unfolding interaction.

For another similar example consider the following:

NP1, Tape 6, 1.29

1 Lot: .t.h The twunny six of Au::gus: A:deline 'er husbin
so we gotta get twenny up.
(0.4)
→ Lot: t. It [ho:lds] fordy seven wir j's gonna ta:ke twenny
5 → Emm: [°u-Oh:°]
(.)
→ Lot: It's twunn[y dollars apiece which isn' ba:d.
→ Emm: [oh
(0.3)
10→ Emm: .t.hhh (.) Oh: maybe he might go:?
(.)
Lot: Hu:h?
Emm: °Wul° 'e mi[ght go] ag[ain.]
Lot: [it's on]na [Mon]dee

In line 4 Lot provides an informing. Emm reacts to the informing by saying *oh* in complete overlap, in a lower volume, and lower in pitch, as described above. Following a pause, Lot continues, first with another informing “it’s twunny dollars apiece” and then with an assessment “which isn’ bad”. Emm’s overlapped *oh* occurs

during the informing part of Lot's turn. Following the turn, Emm takes over the next action, first by repeating the *oh* and then asking "maybe he might go?"

When *oh* occurs within the first turn informing in overlap, then speaker 1 may self select and continue the turn with another informing, as in the informing chains described above. The floor, however, must be negotiated at the end the informing when the *oh* occurs in overlap. In the above example, Speaker 1 continues the next action in line 7. However, speaker 2 continues the next action in line 10.

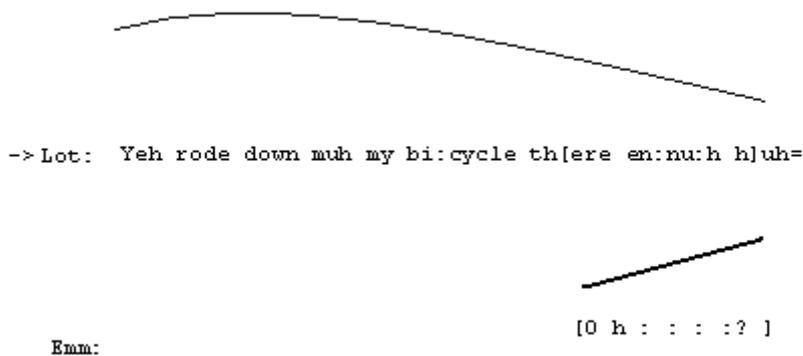
Another similar example may be seen in the following:

NP1, Tape 7, 0.30

1 Emm: Well Bud hadtuh play go:lf uh Thursdee.
(.)
Emm: So'e [didn'take] Sa-uh f-] Fridee o:oo s[o
→ Lot: [O h : : :] :_: : :.] [Yeh rode down
5 muh my bi:cycle th[ere en:nu:h h]uh=
Emm: [O h : : :_:?]
Lot: =wz nobuddy wa(h)s the _:re.
Emm: On Fridee hu[:h?
Lot: [Ye:ah.

In line 3 Emm gives an informing, and Lot offers her *oh* in direct overlap in line 4. Her *oh* is lengthened, and like the nested *ohs* above, it mimics the pitch contour of the informing, i.e. it falls as the informing falls. Like the nested *ohs* described above, it is low in volume and low in pitch. Emm, however, does not continue the turn, rather Lot does. In lines 4 and 5 Lot continues the next action with her own informing "yeh rode down on my bicycle there and ". In complete overlap with this informing is Emm's *oh* in line 6.

At first sight, Emm's *oh* in line 6 seems like a deviant case. Although it occurs in overlap, it does not show the same prosodic stylization as the others. Rather than mimicking the pitch contour of the informing, i.e. falling as the informing falls, the pitch contour rises sharply. In contrast to the low volume and low pitch in other nested *ohs*, it is both high in pitch and high in volume. A visual depiction of the two overlapped pitch contours may be schematically depicted as follows:



Although placed as a nested gestalt, this second *oh* is not accomplishing the same interactional task as Lot's in line 4. Emm's *oh* here is heard as a repair initiator. And indeed, Emm repeats her request for repair in line 7: "on Friday huh?", to which Lot then answers "yeah".

Whereas nested *ohs* in informing sequences are produced low in volume and low in pitch in order to avoid acoustic interference with the unfolding informing, a repair "oh" such as Emm's contribution in line 6 has a different interactional purpose, and might therefore exploit different prosodic practices. Perhaps the steeply rising prosody and high volume contextualizes dissatisfaction with the previous utterance. If so, they would be particularly suitable for initiating repair. For whatever reason, it would be worth another study to look at stylized *ohs* in repair sequences.

The above examples show that nested *ohs* which are low in volume and low in pitch allows the receiver of news to immediately display a change of state as the news is given in the unfolding interaction. I contend that the prosodic practices of low volume and low pitch allow the speaker to communicate this without audibly interfering with the unfolding informing. Indeed, in all the examples in the data corpus the first speaker never seems disturbed by the overlapped *oh* but smoothly continues the informing until the end of the turn. I have shown that following such an informing, where the *oh* is nested within the gestalt of the informing, the initiator of the next action must be renegotiated.

The next section will now look at how the recipient of news can take over the next action following a news informing.

4 Complex oh gestalts

Whereas the first section showed speaker one continuing the next action after an *oh* which was not in overlap and the second section showed the renegotiation of the floor following a fully nested *oh*, the next section will show how second speakers take over the floor by coordinating the placement of their contributions as complex gestalts within the first turn informing. In chapter 1 (p. 54) complex gestalts are defined as “those gestalts in which two gestalts, A and B, form a pattern such that B starts after A starts but before A ends and ends after A ends”.

In the following group of complex gestalts the *oh* following the informing begins before the informing is completed. For this group I have found **no** free standing *ohs*.

One explanation is that the longer turns are used to take over the floor and provide an opportunity for the second speaker to introduce the next action. A schema of the following would look like this:

Speaker 1: informing
Speaker 2: *oh* (in overlap) new action

Prosodically placing the *oh* in this way, and stylizing it with high volume and high pitch is heard as a turn competitive incoming (French/Local 1983). Such a turn competitive incoming can be seen in the following:

NP1, Tape 9, 1.25

1 Emm: So we didn't stay fer th'dinner we jis (.) kahnda l*e:ft
'n came [h o m e]
→ Lot: [Oh: ye]:ah=
→ Lot: =That's 'at a:nnual thing evry[year.]
5 Emm: [iYe::]ah,
(0.3)
Lot: °Ye:[:ah. °]
Emm: [You be]en fi:shin?

In line 1 Emm offers an informing. Towards the end of that action Lot, in overlap, says *oh yeah*, and then, in contrast to the examples described in the first section, continues the next action: “that’s at the annual thing every year”. Lot’s contribution is not prosodically stylized: i.e. the pitches on “oh yeah” are not heard as creating an interval. Prosodically the turn resembles the kinds of *ohs* with additional turn components which are not in overlap with one exception: they are started before the informing end. This can be seen as an additional practice for the recipient of an informing to secure the floor (see French/Local 1983).

High volume and starting the turn before the next turn is over could be a common way for interactants to take over the floor in activities other than informings. In informing activities, however, interactants specifically react to the informing with their “oh”, but the additional turn components added to the *oh* turn make up the new action.

Consider another example of a complex *oh* turn:

NP1, Tape 17, 5.18

1 Mar: [.hHHHH En then'e has uh: (.) this uh College Pa:rk which is
up by Fullert'n it'[s,
Edn: [.hHHHH[hHHH]
Mar: [it's] [c'nected with] the]
→ Edn: [OOoo::that's L]Y::]NCH!
(0.2)
Mar: Ye:ahh. 5:30
Edn: Gad it wz in the pa:per[::
Mar: [Mhm w'l tha:t's Larry's story.
10 (0.5)
Edn: Is that LERRYISS?
Mar: 'At wz Larry's story yeah. [He wrote it,]

Unlike in the above example, Edn, following Mar’s informing in lines 1,2 and 4 heavily stylizes her *oh* contribution. Indeed, her *oh* is so loud and lengthened that Mar is forced to abandon her turn in the middle and orient towards Mar’s *oh* turn, which she does in line 7, “yeah”.

NP1, Tape 10, 0.20

1 Emm: [n:No::: I jist hadda lo:c'l
dea:l en:nuh id wud'n any fu:n but I'm BETTER I wz: lying

on: the cou:ch out'n [front]
 → Nan: [e-Oh:]:: I:'m sorr[y E:m]ma:h,hh
 5 Emm: [AH:..]
 Emm: ((cutesy)) dI AM TOO WHYNCHE COME'N SEE m[e.

I contend that recipients, by placing their *oh* turns as complex gestalts in relation to the first turn informing, are able to take over the floor. These are, as Local has shown, typical competitive floor incomings. With informing sequences, however, a second speaker must make her claim to the floor particularly explicitly, as *oh*, in most other cases, is simply used to show a state change, but does not, in itself, demand a reaction to it by the other interactant. By itself, *oh* or *oh* plus an assessment does not demand the other interactant to respond. When the recipient of news does desire to take over the floor, then she must do decisively, which is why a complex gestalt and prosodic stylization (high volume and lengthening) seem particularly suitable for accomplishing the task.

5 Conclusion

This chapter has shown how the practices of prosodic stylization and gestalt placement influence turn taking and floor holding in informing sequences. The chapter has shown three ways in which *oh* is placed in relation to the informing and the influence this and the prosodic stylisation of *ohs* has on the unfolding interaction. It has been shown that free standing *ohs* which occur in the free following a first turn informing and which are realized with low volume and a falling pitch are followed by speaker one taking the initiative and continuing with the next action. Non-overlapped *ohs* which occur with an assessment are realized with a harmonic interval. When such *ohs* are not stylized, then the second speaker is heard to be making claims to the floor. In such cases it is the news recipient who precedes to initiate the next action.

*Oh*s which are realized as nested gestalts and which show low volume and low pitch are used in this way in order to show change of state immediately as the informing unfolds. Such *ohs* are heard to mimic the falling pitch contour of the informing turn in which it is nested. I claim that the low volume and low pitch of the *oh* allow the

speaker to show her change of state without disrupting the first turn informing with audible interference.

Finally, in complex gestalts, second speakers claim the floor for the next action by placing their contributions towards the end of the first turn informing and continuing after the informing has come to an end. Such *oh* turns may be heard as *ohs* with additional turn components which are not in overlap (and therefore lack prosodic stylization), or they may be realized with strong stylization: lengthening, high pitch and high volume. I argue that such stylization allows the second speaker to competitively come in and make claims to the floor.

This chapter has also discussed some deviant cases in which the *oh* has been shown to display other gestalt placements and prosodic stylizations. In these cases, however, it was shown that the interactants were engaged in an activity other than an informing sequence. Another study could take a similar approach to these same practices in such activities as repair, for example.

Chapter 5: Prosodic Stylization In Assessment Sequences

The following intends to expand ongoing research into the practices used for expressing preference within assessment activities. Pomerantz's (1975, 1984) initial observation that preference is expressed lexically as well as through pause and overlap has been refined by later research to include other practices such as the use of particles (Auer/Uhmann 1982), the use of overlap (Goodwin/Goodwin 1987, 1992) and the use of speech rhythm (Uhmann 1996; Auer/Couper-Kuhlen/Müller, 1999). My observation is that not only is interactional rhythm used as an "independent gloss" for supporting preference structures (as argued by Auer/Couper-Kuhlen/Müller, 1999), but prosodic highlighting and stylization in general can be used for this purpose within assessment sequences. Furthermore, highlighted and stylized prosody may also indicate upgrades and downgrades in second and third turn assessments even when, lexically, the contribution involves the same term. I will examine two kinds of assessment contexts in which prosodic stylization plays an important role in 1) establishing rapport via co-participation in the activity and 2) contextualizing emphasis³² and cueing a playful modality.

1 Introduction

In spoken discourse, when does an utterance (or part of an utterance) draw attention to itself and become stylized? This question may be answered in a number of ways, depending on the linguistic level or levels which one is considering (syntactic, semantic, pragmatic, etc.). Considering the prosodic level, I propose a gestalt approach for asserting when a segment of speech becomes stylized.

The ability of stylized prosodic structures to contextualize intensity and emphasis in interaction has developed certain functions within particular kinds of speech activities. One sees this clearly in assessment sequences, in which prosodic stylization can be created collaboratively to heighten the involvement and rapport between speakers as

³² Susanne Uhmann shows that beat clashes contextualize intensity and emphasis (1996, 358-359). I assert that not only beat clashes (a kind of prosodic stylization) contextualize intensity and emphasis, but so does prosodic stylization in general.

they co-participate in the making of assessments and displaying of affect. Additionally, one sees this when participants use prosodic stylization in first and second assessment sequences, thereby ignoring the risk of a dispreferred next turn. In both cases, the enhanced prosody plays an important role in separating the marked speech from what has come before.

Assessing something is a delicate activity and, depending how the speaker forms his/her turn, frequently has the potential to evoke a dispreferred next turn. As Auer/Uhmann have pointed out, offering a first assessment, particularly when one does not know the other participant well, can be risky:

Der Produzent einer ersten Bewertung ist dem der zweiten Bewertung gegenüber in der 'schlechteren' Position.

(1985:5)

The speaker of a first assessment is in a worse position than the speaker of a second assessment.

(my translation)

In the following I will argue that in certain assessment contexts interactants do, in fact, take this risk. Participants form their prosody regardless of dispreferred next turn shapes: for the fun of it. In playful modalities, I argue that participants exploit the emphatic aspects of stylized prosody in assessments for the purpose of provoking responses which are creative on a number of levels: including prosody, lexis, syntax, etc.

Additionally to the upgrading and downgrading of lexical terms as shown by Pomerantz (1975, 1984), the use of particles as shown by Auer/Uhmann (1982), and the use of pause, overlap and other rhythmic phenomena to cue preferred or dispreferred responses as shown by Goodwin/Goodwin (1986, 1992), Uhmann (1992), Auer/Couper-Kuhlen/Müller (1999), the highlighting and stylization of prosody within assessment sequences also allows conversation participants to align or nonalign themselves with the other participants, thus allowing them to manage a potentially difficult social activity. Within the second turn of assessment sequences prosodic stylization may support lexical upgrades and downgrades, but it may also

give the impression of an upgrade or a downgrade even when the lexical evaluation within the second assessment is similar or the same as the initial assessment.

The following is divided into three parts. In the first part I will give an overview of some of the work that has been done on assessments from a talk-in-interaction point of view. This overview begins with the groundbreaking work of Pomerantz (1975, 1984) and her findings with respect to preferred and dispreferred responses in upgraded, downgraded and same evaluated second assessments. Since her initial work, many researchers have shown that beyond lexis and pause/overlap, other considerations are important for how preference is signaled in discourse. Following the overview of Pomerantz, I will consider Auer/Uhmann (1982) Goodwin/Goodwin (1986/1992), Uhmann (1996) and, most recently, Auer, Couper-Kuhlen and Müller (1999). In the second part I will offer an example of what I mean by prosodic stylization, using an example of an assessment. I will show that upgrades and downgrades may also be signaled by prosodic stylization. Finally, I will develop my assertion that prosodic stylization is an important practice in co-participating in making assessments and displaying affect.

2 Previous Work on Assessments

The following section will give an overview of previous talk-in-interaction research on the activity of assessment. I will start with Anita Pomerantz's observation that second assessments may be formatted as preferred or dispreferred depending on certain conversational practices such as upgrading, downgrading or "same-grading" and the systematic use of pause and overlap.

Following the overview of Pomerantz's work, I will discuss other scholars who have refined her findings by more closely considering other practices important for preference/dispreference including the use of particles in German and the use of pause and rhythm. A discussion of the work of Auer/Uhmann (1982), Goodwin/Goodwin (1987, 1992), Uhmann (1996) and Auer, Couper-Kuhlen, Müller (1999) will be the focus of the second part of this section.

2.1 Pomerantz (1975, 1984)

Some of the earliest and most influential work on assessment sequences are Anita Pomerantz's 1975 PhD thesis Second Assessments. A study of some features of agreements/disagreements (University of California at Irvine) as well as her 1984 article "Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes". Here Pomerantz argues that following an initial assessment, second speakers proffer a second assessment which shows their agreement or disagreement with the assessment via preferred or dispreferred next actions. Such next actions may take the form of an upgrade, a same evaluation, or a downgrade. Upgrades, according to Pomerantz, "may be considered strong agreements on sequential grounds". Upgrades commonly have two formats. In the first, the second assessment may have a stronger evaluative term than in the first assessment. An example can be seen in the following example:

2.1.1 (JS:II:28)

J: T's tsuh beautiful day out isn't it?
L: Yeh it's just gorgeous...
(1984:65)

In this example the second speaker, L, upgrades J' initial assessment "beautiful" to "gorgeous". In the second format the speaker may add an intensifying modifier to the same evaluative term as in the initial assessment:

2.1.2 (CH:4.-14)

M: You must admit it was fun the night we
we[nt down
J: [It was great fun
(1984:66)

In this example, J adds the modifier "great" to the same evaluative term, "fun", that M uses in the initial assessment. According to Pomerantz, these kinds of upgraded

second turns following initial assessments are considered “strong agreements on sequential grounds” (1984:66).

Second speakers may also proffer the same evaluation as in the initial assessment, which may either indicate agreement or preface disagreement depending on the context and what follows. This first example indicates agreement via same evaluation:

2.1.3 (JK:3)

C: ... She was a nice lady - I liked her
G: I liked her too

(1984: 67)

Here, the same evaluation in the second assessment prefaces disagreement:

2.1.4 (J&J)

A: Yeah I like it ()
B: I like it too but uhh
Hahheh it blows my mind.

(1984:67)

Pomerantz claims that same evaluations before disagreements nonetheless “may be considered a kind of weak agreement” (68).

Frequently, according to Pomerantz, downgraded sequences engender disagreements. In the examples she gives, the downgrading sequences all involve choosing a downgraded evaluative term, such as in the following.

2.1.5 (GJ:1)

A: She's a fox
B: Yeh, she's a pretty girl
A: Oh, she's gorgeous!

(1984:68)

In this example, the evaluative term “fox” in the initial assessment is downgraded to “pretty” in the second assessment. Following such downgrades, one response is for the first speaker to restate his/her initial position, frequently in an even stronger way than in the initial assessment. In the above example, following the downgraded assessment “pretty girl”, A chooses the stronger evaluative term “gorgeous” in her turn.

Not only is lexis important for establishing agreement, so is prosody. Pomerantz fleetingly discusses the role of prosody within assessment activities, and asserts that the pauses and overlap at the turn transition place indicate whether the next action is preferred or dispreferred. Preferred next actions will be connected to the initial utterance with a minimization of gap, or may even occur in slight overlap. This is displayed in the following examples:

2.1.6 (NB:PT:19:r)

L: God it's good.=
E: = Isn't that exci:ting,

(1984:69)

2.1.7 (SBL:2.1.8.-5)

B: She seems like a nice little [lady
A: [Awfully nice
little person.

(1984:67)

Dispreferred next actions may display a pause, or even silence. Pomerantz has observed that “incorporating delay devices constitutes a typical turn shape for disagreements when agreements are invited” (70). This can be observed in the following example:

2.1.8 (NB:IV:11.-1)

A: God izn it dreary
(0,6)
Y'know I don't think-
B: .hh It's warm though,

(1984:70)

Pomerantz can be credited with uncovering and describing the sequential organization of assessment sequences as well as some of the practices used to accomplish this activity. I will now turn to some of the scholars who have refined Pomerantz's initial observations in order to show that other practices besides lexis, pause and overlap as described by Pomerantz play a considerable role in assessment sequences.

2.2 Auer/Uhmann (1982)

Auer/Uhmann (1982) apply Pomerantz's analysis to spoken German data, and find that some revisions are in order. In contrast to Pomerantz's assertion that same-graded lexical evaluations may indicate either agreement or disagreement depending on the following utterance, Auer/Uhmann show that such same-grades only preface disagreement. Furthermore, in spoken German data, particles already indicate agreement and disagreement even before lexical upgrading/downgrading/same-grading takes place.

The importance of Auer/Uhmann for the following is in showing that there are, in fact, other practices beyond lexical upgrading, downgrading, same-grading and beyond pause and overlap which demand consideration when analyzing assessment sequences.

2.3 Goodwin/Goodwin (1987, 1992)

Pomerantz's observation that agreements frequently occur in slight overlap is further taken up by Goodwin/Goodwin. Although they do not directly deal with the notion of preference and dispreference as described by Pomerantz, Goodwin/Goodwin do show that overlap – which is described by Pomerantz as being an action that is preferred within assessment sequences – often displays “a state of heightened mutual involvement” (1992, 79). Such overlap, often coupled with non-verbal cues such as head-nodding, “constitute a very strong way of doing agreement” (1992, 78). Goodwin/Goodwin, noting that the sequential positioning of assessments is frequently found in the *middle* of turn-construction units, assert that first speakers may project an oncoming assessment and that second speakers may overlap this assessment, thereby *co-participating* in the assessment. The following example exemplifies co-participation as described by Goodwin/Goodwin:

2.3.1 “Asparagus Pie”

Nancy: *Jeff* made an asparagus pie
 It was s::so [:goo:d
Tasha: [I love it. Yeah I love that

(Goodwin/Goodwin1992:81)

Goodwin/Goodwin argue that Nancy projects her assessment with the intensifier “so”, spoken with what they call “enhanced intonation”. By projecting her own assessment, “good”, in this way, Nancy allows Tasha the opportunity to come in simultaneously with her own assessment, “I love it”. This allows the co-production of the assessment by both speakers in what Goodwin/Goodwin call a “climax” or “peak” (81).

As in the Auer/Uhmann study, Goodwin/Goodwin show that there are other practices involved in assessment sequences that must be considered. They show that prosody is used in a complex way to organize discourse. Goodwin/Goodwin assert that assessments may become a collaborative, interactive activity in which participants display heightened mutual involvement.

2.4 Uhmann (1996)

Susanne Uhmann specifically links a certain prosodic practice with preference structures in second assessments in her article “On rhythm in everyday German conversation: Beat-clashes in assessment utterances” (1996). Uhmann defines beat-clashes as “highly marked rhythmical structures in which the phonologically unmarked alternation between prominent and non-prominent syllables is cancelled in favour of a succession of prominent syllables”. An example of a beat-clash can be heard in the following example:

2.4.1 “On this Issue Alone”

On the Barbara Carlson and Friends Show, radio talk-show moderator Barbara Carlson talks about voting for a mayoral candidate on the basis of the single issue of abortion.

```
/is /  
/it /  
/the /  
/big /  
/deal /  
/(.) I/  
/mean  
are you gonna vote for these people  
on this issue alone
```

In the first five lines, each syllable receives equal stress. Uhmann asserts that such beat-clashes are restricted in their sequential occurrence in assessment sequences. According to Uhmann, “beat-clashes occur in extended first assessments like stories, news or informings and in seconds to these conversational objects, but they are absent in first and second assessments of assessment pairs” (303).

Uhmann proposes an explanation for why beat-clashes do not occur within initial and second assessments in her data. She claims that “the absence of beat-clashes in assessment pairs is not accidental but systematic (1996:258).” Noting Auer/Uhmann’s (1982) findings that second assessments must display upscaling in order to display agreement, Uhmann asserts that a second assessor is unable to

upgrade when the initial assessment is “near the end of the evaluation scale” (353).

According to Uhmann:

If a first assessor presents an assessment that is further strengthened by a beat-clashing rhythm, the second assessor faces the problem that proffering an unambiguous agreeing second assessment implies the production of an assessment that has to be at least as strong as, or even stronger than, the first. How can this be achieved? It can only be achieved by presenting a stronger or equally strong assessment term realized with a beat-clashing rhythm. This is of course not impossible (...), but as assessors are held responsible for the positions they state, first assessors systematically increase the possibility of stating a first assessment that cannot be agreed with if they produce an evaluation which is near the end of the evaluation scale.

(Uhmann 1996:353)

Uhmann asserts that beat-clashes contextualize “intensity” or “emphasis” (358) and thus would be unlikely to occur in assessments. Uhmann says a first speaker risks a dispreferred next action after an initial assessment which has been realized with a beat-clash. According to Uhmann: “as agreeing by producing an upgraded or at least same assessment is preferred, interlocutors have to orient to this preference so that assessment pairs that show this formal preference can be achieved” (1996:353). Participant’s avoidance of beat-clashes in first and second turns leads Uhmann to conclude that such avoidance shows “participant orientation towards a formal preference for agreement” (1996:359).

Uhmann’s observation is directly relevant to my analysis of stylized prosodic gestalts in assessment activities. Beat-clashes are, in my sense, a form of prosodic stylization. The further role of such stylized prosodic gestalts within assessment sequences will be taken up in the following sections.

2.5 Auer/Couper-Kuhlen/Müller (1999)

Most recently, Auer/Couper-Kuhlen/Müller have taken up the notion of preference in the chapter “Rhythm and preference organization in English” from their book Language in Time (1999). They deal with a variety of activities, including among others, informings, repair, as well as assessments and pointedly ask, “can marked

rhythm be regarded as a sign of dispreference”. Whereas Pomerantz would answer “yes” on formal grounds and Atkinson and Heritage would answer “yes” on substantive grounds, Auer/Couper-Kuhlen/Müller answer “sort of”. In their own words, they assert: “interactional rhythm does not correlate neatly with either a formal or a substantive definition of preference but constitutes instead an independent ‘gloss’ which interacts in predictable ways with preference options” (92).

In the first section of the chapter, Auer/Couper-Kuhlen/Müller contradict previous assertions about timing and preference by asserting that “rhythmically integrated timing does not correlate consistently with what is canonically considered to be a preferred action” (102). By presenting examples of rhythmically integrated disagreement and rhythmically non-integrated agreement sequences, they show that there is no formal connection between rhythm and preference.

Auer/Couper-Kuhlen/Müller then proceed to look at preference as a substantive notion as proposed by Atkinson and Heritage (1984), who suggest that “the design features associated with the production of preferred/dispreferred activities may inform and be informed by a logic of ‘face’ considerations (Brown and Levinson 1978, Goffmann 1955) at the levels of both form and usage” (1984:54. In: Auer/Couper-Kuhlen/Müller 1999:102). According to Auer/Couper-Kuhlen/Müller “the crucial claim in a functional approach to preference is to think of face as determining what is regarded as a preferred or dispreferred course of action” (1999:102). In other words – in a general way – preferred actions maintain face and dispreferred actions threaten face. With a substantive notion of preference, unmarked rhythmically integrated second assessments will be associated with face-saving/face preserving; marked rhythm will be associated with face-threat.

Although such a notion of preference can explain more cases than the formal notion as described above, it does not explain all cases. Auer/Couper-Kuhlen/Müller proceed to consider so-called “hot-news informings”, in which the substantive notion of preference does not hold. Silent beats and late stresses, contextualizing unexpectedness and recognition that the news is indeed “hot”, are the preferred practices after such news informings. According to Auer/Couper-Kuhlen/Müller, “hot-

news informings with rhythmically delayed but unmarked responses, appear to weaken the claim (...) that rhythmic integration is the unmarked alternative”.

What role does rhythm then play in assessments? Auer/Couper-Kuhlen/Müller claim it is an “independent gloss of preference” (107). They assert “interactional rhythm is an independent glossing device which takes its primary signalling value from its use in congruence with preference status” (115). In other words, in some contexts rhythmic integration will be preferred, in others dispreferred; in some contexts rhythmic non-integration will be preferred, in others dispreferred. Auer/Couper-Kuhlen/Müller claim that “both marked and unmarked rhythm then ‘gloss’ situated actions - whether preferred or dispreferred – in characteristic ways” (115).

What the above discussion of past research has shown is that Pomerantz’s initial findings, as important as they were, are open to refinement and improvement. Not only are more practices involved in preference and dispreference structures than Pomerantz indicates, but the notion of preference itself and the role of prosody for displaying preference have been criticized. It has clearly come out in later research that the prosodic element of timing is a crucially important aspect in assessment sequences. Prosody, however, consists of more than just timing; it includes pitch and volume as well. In the next section I will describe prosodic stylization using an example from our corpus of spoken data. Following this, I will, in the final section, explain why pitch and volume as well as rhythmic stylization are necessary considerations for the activity of assessing.

3 Upgrades/Downgrades through Prosodic Stylization

What else prosody, or more specifically prosodic stylization, might be doing within the activity may be indicated in two of Pomerantz’s own examples of second assessments. The first is an upgraded assessment:

3.1 (MC:1)

A: Isn’t he cute
B: O::H he::s a::DORable

(1984:65)

The second is a downgraded second assessment:

3.2 (GJ:1)

A: She's a fox!

B: Yeh, she's a pretty girl.

(1984:68)

In the first example, the extreme lengthenings on “O::H”, “he::s” and “a::DORable” and the exclamation mark in the second example following “fox” certainly indicate that some kind of heightened prosodic design is involved. Although I don't propose a formal correspondence between prosody and preference in assessments, it does appear that prosodic stylization is underscoring the lexical upgrade in the first example, and the return to “normal” after the heightened intonation of the first assessment underscores the lexical “downgrade” in the second example.

For an example of stylized prosodic gestalts, consider the following sequence, taken from the radio talk-show program, *The Barbara Carlson and Friends Show*:

3.3 “Fabulous”

1 B: .hh some of my friends
2 Peter Theo
3 are choosing (.) to live (.) with men
4 rather than marry them
5 P: are these older babes
6 B: these are friends
7 P: middle age babes
8 B: would you call nurse Ratchet an older babe (.) Peter
9 P: yeah from haha my from my perspe(h)ctive(h) yes
10 B: she looks damn good though doesn't she
11 P: she's in good shape
12 B: well she's in good shape because she's had=
=everything done
13 I'm mean she's got died hair
14 → she's false boobs false hair false eyes
15 → I mean she is making herself look fabulous
16 (.)

17 that's probably why she has somebody to live with=
=her

Starting in line 12 Barbara begins to stylize her prosody, at first moderately, and then progressively more until the final line “I mean she is making herself look fabulous” is heard as a prosodic *tour de force*. This final line incorporates a variety of prosodic manipulations which make it the climax of an increasingly stylized sequence.

Barbara’s contribution from lines 12-15 displays two isochronous rhythmic gestalts. The first can be heard in line 12-13:

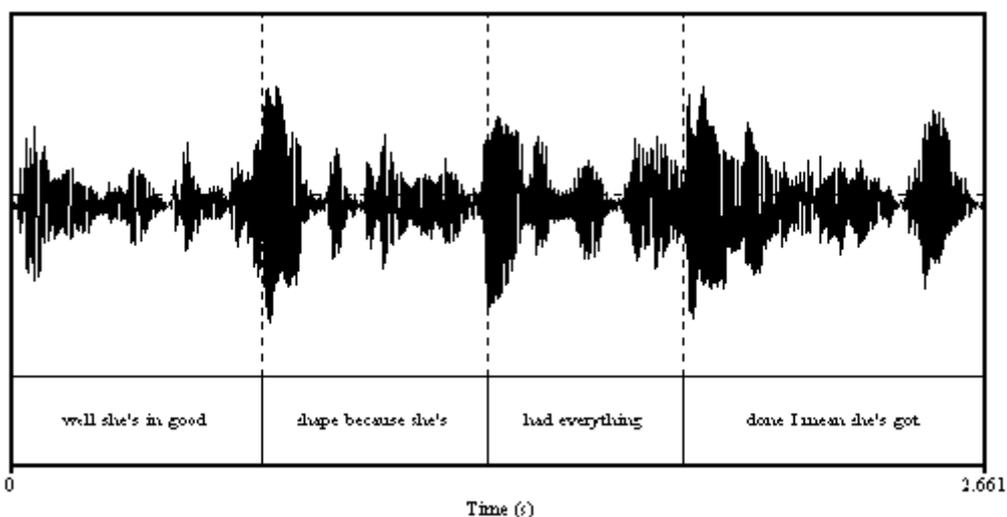
3.4

/well she's in good /
/shape because she's /
/had everything /
/done I mean she's

In these lines, the stress on the main syllables “well”, “shape”, “had” and “done” fall at equal distance from one another. Barbara herself further displays a heightened orientation to the rhythm of her contribution by simultaneously tapping to the syllables “shape” and “done”.

The following wave form supports our auditory perception acoustically:

3.5



The acoustic wave shows that the points marking greatest intensity fall isochronously. The distance between the highest intensity of “well” and the highest of “shape” is 0.69 seconds, between “shape” and “had” is 0.63 seconds, and between “had” and “done” is 0.6 seconds.

The second rhythmic gestalt occurs immediately following the first. Barbara “jumps the beat” in line thirteen and then produces a second rhythmic gestalt that increasingly becomes more stylized:

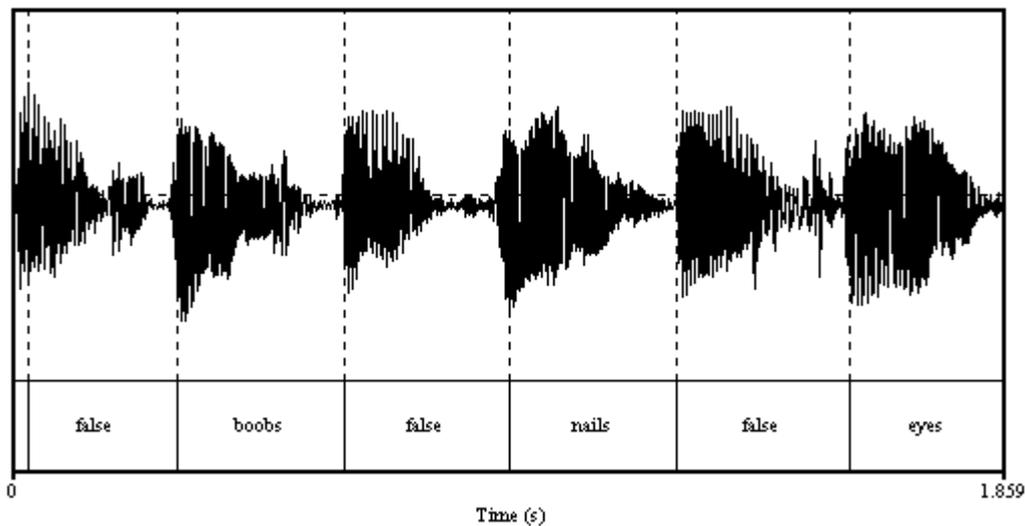
3.6

```
          I mean she's got  
/died hair she's /  
/false ↑boobs  /  
/false ↑nails  /  
/false ↑eyes
```

In this example each syllable - “false”, “boobs”, “false”, “nails”, “false” and “eyes” – occurs at isochronous intervals. Pitch step-ups additionally bring out the syllables “boobs”, “nails” and “eyes”, making the rhythm isometric in Auer/Couper-Kuhlen/Müller’s sense (1999), who define isometric rhythm as rhythmic intervals in which “in addition to being perceived as temporally regular; contain an equal number of syllables” (1999:40). The isometric and isochronous rhythmic gestalt is further brought out by Barbara through her accompanying rhythmic tapping.

As for the above stylized rhythmic gestalt, an acoustic wave form is presented here in order to show the rhythmic gestalt which we perceive auditorily:

3.7



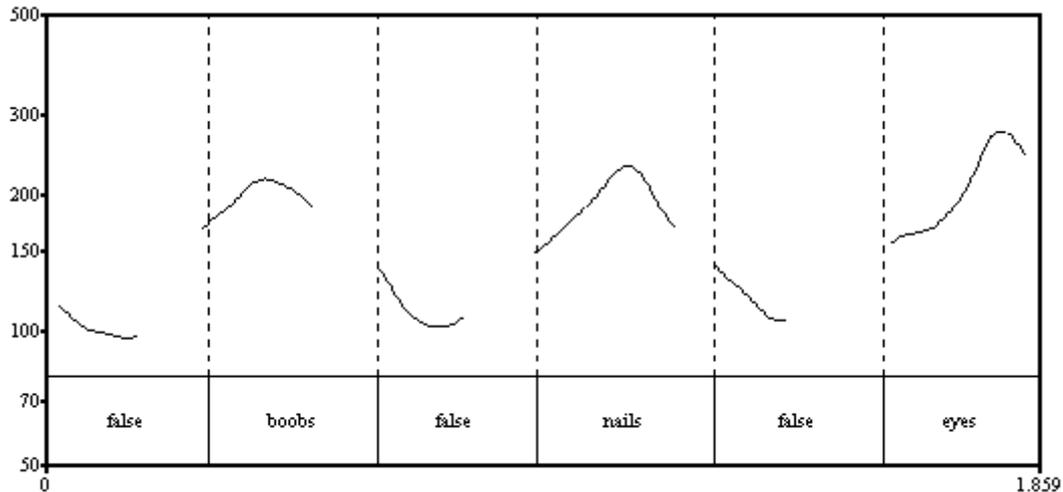
Here the intervals between the points of greatest intensity are .31 seconds between “false” and “boobs”, .32 seconds between “boobs” and “false”, .31 seconds between “false” and “nails”, .32 seconds between “nails” and “false”, and again .32 seconds between “false” and “eyes”.

This excerpt not only shows a strong rhythmic pattern, a stylized pitch pattern also may be heard. Consider again line 14:

she's false boobs false nails false eyes

In this excerpt, Carson can be heard to orient the pitch of her contribution towards two pitches: a lower one and a higher one. Whereas the upper pitch gets higher, the lower pitch stays the same. I propose that this increasing of pitch in the top line but not in the bottom line increases the intensity of the emerging utterance. This is supported by Barbara's tapping. Barbara rhythmically taps while speaking. At the climax of the rising pitch pattern on “eyes”, she not only taps, but also can be heard to slap her hand down on her desk. Barbara highly orients towards one lower and one higher pitch. Such a perception may – like the rhythmic gestalts above – be supported through acoustic analysis; this time, however, through a frequency analysis:

3.8



The frequency analysis, presented logarithmically to match our perception of pitch to fundamental frequency, shows that the pitch of the bottom line at the syllables “false”, “false” and “false” stays constant at approx: 100 Hertz. The peaks, however, on the syllables “boobs”, “nails” and “eyes” consistently rise. Moreover, they rise in a patterned way.

All of this stylization culminates in the following line, in which Barbara - within a few seconds of her initial assessment and Peter’s second assessment - offers yet a new assessment:

she’s making herself look f::a:::bulous

Prosodically, this intonation phrase is realized generally higher than her previous utterances. The extreme syllable lengthening brings out the word “fabulous”, which, lexically, is the evaluative word within the assessment. “Fabulous” is additionally realized with more energy than the surrounding discourse and is perceived as louder.

The example shows that not only lexis is important, so is prosody. The lexical difference between Barbara’s first assessment “damn good” and her later assessment “fabulous” is more or less equal. “Fabulous” cannot be heard either as a

lexical upgrade or a downgrade with respect to “damn good”. Considering the highly stylized prosody, however, one most certainly gets the impression of an upgrade.

Why did I choose the above example? For one, it contains three assessments: 1) “she looks damn good” 2) “she’s in good shape” 3) “she’s making herself look fabulous”. Secondly, the excerpt is amusing - which is an important consideration for radio and television broadcasters, particularly of the “talk-show” ilk. Finally, beyond its “aesthetic” character, the prosodic stylization is crucially important for the talk at hand and the activity of assessing. It is this third reason which I will explore in more detail in the following section.

4 Prosodic Stylization in Assessment Activities

The following section looks at assessment activities in which I assert prosodic stylization plays an important role, first in establishing rapport and involvement between speakers, and second as a means for cueing a playful modality. This chapter intends to expand Uhmman’s (1996) findings by showing how in some contexts speakers do highly pattern their prosodic contributions within initial and second assessments. Such stylized contributions, formatted collaboratively by a number of interactants, increase the involvement among participants. In other contexts, speakers knowingly – even intentionally – risk a dispreferred next turn by formatting the prosody in a stylized way.

4.1 Collaborative Prosodic Stylization in Assessments

As mentioned in the introduction, assessments frequently have the potential to be a delicate activity. The first assessor presents an evaluation and elicits a second evaluation without knowing for sure whether the second evaluation is necessarily congruent to his or her own assessment. It is therefore risky to offer a first assessment: you offer your cheek without knowing 100% whether you might get kissed or slapped.

Whereas such risk might well be a consideration in conversation among acquaintances or strangers, there is less risk when, for example, one knows one's fellow conversation participants quite well, especially when they are family, close friends, intimates, etc. In some contexts and with such partners one can already judge with reasonable accuracy how they will respond to the initial assessment. There is also a lower risk when the participants are engaged in an activity in which positive evaluations are the established and "only" possible assessments which can be offered. In such cases, the risk of a dispreferred second turn is less likely, and the assessing may be used to establish rapport with the other conversation participants, who themselves give the second assessment at least as much stylization as in the first assessment. For obvious reasons, prosodic stylization with its ability to contextualize emphasis and intensity is an effective practice for establishing heightened involvement among participants.

Consider, for example, the following example, in which Kendra has just opened a birthday present from her brother and sister-in-law:

4.1.1 "Appease The Monster"

In this excerpt, Kendra is opening birthday presents. Five family members are present: Kendra's mother "Marci", her father "Ken", her brother "Kevin", and her sister-in-law "Wendy". The present is from Kevin and Wendy.

1 KENDRA: uh [ʔoh:
2 KEN: [oh: yeah
3 KENDRA: were talking
4 baking mon[ster we're talking
5 KEVIN: [tollhouse morsels
6 MARCI: [definitely more than a pair and a=
7 =spare
8 KENDRA: cookie monster
9 appease the monster
10 h[o::ly cow:::
11 Wendy: [oh::[::
12 KEVIN: [yay:
13 [holy cow
14 KENDRA: [cookie baking se:t
15 MARCI: [a:right
16 a:right
17 KENDRA: mm::[::
18 KEVIN: [rubber [mai:d
19 MARCI: [oh:
20 [let me see it
21 KEVIN: [you can't squash it

22 (.)
 23 KENDRA: mm::::[rubber mai::d
 24 MARCI: [ho::w neat
 25 KEVIN: twelve [pieces
 26 KENDRA: [yay
 27 KEVIN: <gasp>
 28 KEN: [tha[t's I
 29 MARCI: [oh:[: that's
 30 KENDRA: [wo:[:w
 31 KEVIN: [oh that includes all the teaspoons though

Following Kendra's unwrapping of the present there are three kinds of utterances, most of which are prosodically stylized:

- 1) Evaluating and assessing the present: "alright", "how neat", etc.
- 2) Descriptions of the present: "Rubbermaid", "you can't squash it", etc.
- 3) Prosodically stylized response cries: "oh", "holy cow", "mm", etc.

I argue that in this sequence, prosodic stylization is a practice which allows the participants to co-participate in the unfolding activity, sharing not only mutual assessments, but also mutual emotional responses. Response cries are a way of displaying in the "here and now" one's current state (Goffman 1978). Speakers display an orientation to the prosody of the other participants, thereby allowing them to form collaboratively stylized prosodic gestalts.

Goodwin/Goodwin (1992) describe how conversation participants can co-participate in assessments in the article summarized above. Similarly to the participants in Goodwin/Goodwin's example, "Asparagus Pie", the participants here in "Appease the Monster" use the practice of overlap to display co-participation in the activity. Additionally to overlapping assessments, however, participants cooperatively build prosodic gestalts. The effect is one of sharing not only the mutual assessment, but the mutual emotional response as well.

Consider, for example, the first two lines.

KENDRA: uh [?oh
KEN: [oh yeah

Kendra's first reaction after partially unwrapping the present is the response cry in line 1, "uh ?oh". One hears an orientation to pitch on these two syllables. Kendra's top pitch is heard to be a fourth above the bottom pitch. Ken displays his orientation towards the prosody started by Kendra in his contribution in line 2. His "oh" is exactly in overlap with Kendra's "oh". Not only is this "oh" heard to line up with Kendra's "oh", displaying a heightened orientation to the timing of the unfolding utterance, his "oh" is even heard to be in harmony being a fourth above Kendra, which is Kendra's original first pitch. Although the two participants do not "sing" their contributions like a musician, holding pitches, one can still use music notation to present how Kendra and Ken are orienting towards rhythm and pitch:

4.1.2



Kendra: uh oh
Ken: oh yeah

Throughout the sequence various participants orient towards pitch on a single lengthened syllable. The speaker hits this pitch, and then slowly glides down, giving the syllable a distinct pitch contour. This can be clearly heard in the following turns by Kendra and Marci:

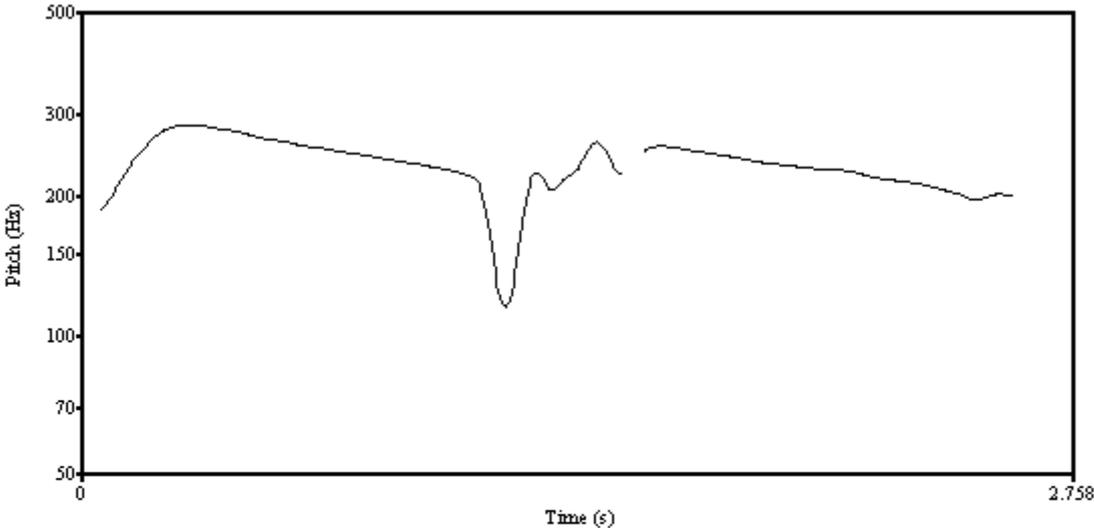
KENDRA: mm::::
 [rubber mai::::d
MARCI: [ho::w nea::t

Kendra starts her turn with a response cry, "mm", which is lengthened, and holds a distinct pitch which glides down. Then there is a pitch step up, and she says "rubber

maid”. The anacrustic syllables “rubber” hit the same pitch as the beginning pitch of “mm”, “maid” starts at this pitch and like “mm” glides down in an identical melodic contour. Marci displays a heightened orientation toward timing and pitch in her assessment, which occurs simultaneously with Kendra’s “rubber maid” and which is also realized at the same pitch and with the same slow downwards gliding. Furthermore, Kendra’s and Marci’s utterances end simultaneously.

The following f0 analysis visually depicts the contours which one perceives:

4.1.3



Kendra:	mmm:::	rubber	maid
Marci:		how	next

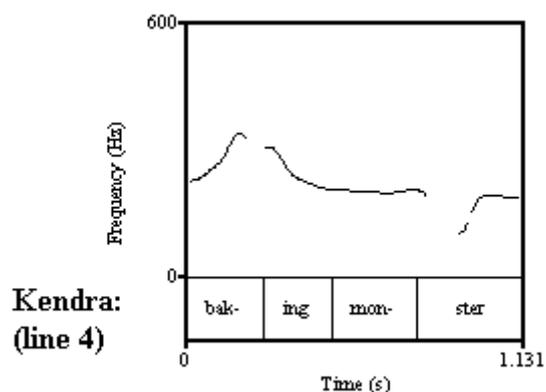
In these utterances there is an orientation towards a distinct pitch, seen to be at approx. 300 Hertz in the above frequency analysis. The pitch glides are approx. 0.9 seconds long. The pitch glide on both contours is easily recognized in the above graphical representation, both ending at approx. 200 Hertz.

This pitch contour is deployed by other participants as well. Following is a list of fifteen further objects which display a pitch contour similar to the one described above:

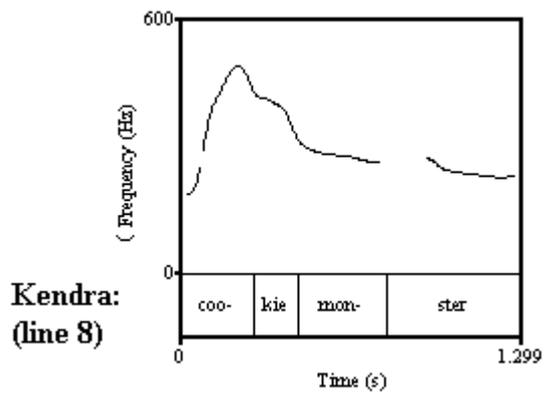
- 1) baking monster (line 4)
- 2) tollhouse morsels (line 5)
- 3) cookie monster (line 8)
- 4) appease the monster (line 9)
- 5) holy cow (line 10)
- 6) oh (line 11)
- 7) yay (line 12)
- 8) alright (line 15)
- 9) alright (line 16)
- 10) mm (line 17)
- 11) rubber maid (line 18)
- 12) mm (line 23)
- 13) how neat (line 24)
- 14) oh (line 29)
- 15) wow (line 30)

Those contours which do not occur in overlap and are therefore conducive to a frequency analysis are depicted below:

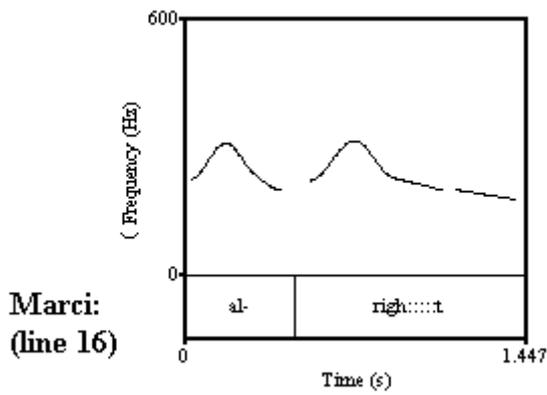
4.1.4



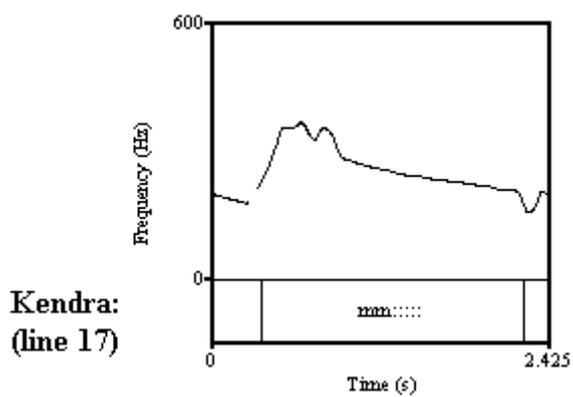
4.1.5



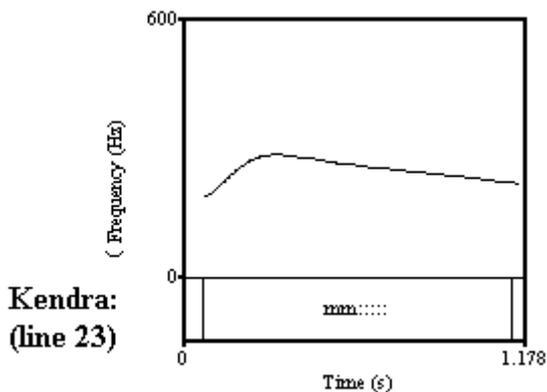
4.1.6



4.1.7



4.1.8



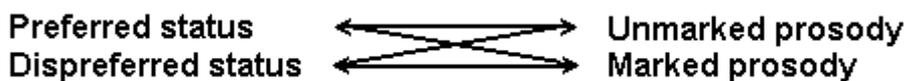
The five f0 analyses above show that the pitch configuration of the contributions starts at approx. 300 Hertz and falls to approx. 200 Hertz. The gliding contour is in most of the examples approx. 0.9 seconds long. Furthermore, the gestalt pattern is produced by a number of different speakers: Kendra, Kevin, Wendy and Marci.

Many of these utterances are also in overlap, creating a strong sense of assonance. Such overlap coupled with an orientation to pitch can be heard in lines 1 and 2, in lines 4, 5 and 6; in lines 10, 11, and 12; in lines 13, 14 and 15; in lines 23 and 24, and, following Kevin's gasp in line 27, in lines 28, 29 and 30.

The above excerpt shows how a group of participants can not only display co-participation in assessments via overlap, as asserted by Goodwin/Goodwin; they have an arsenal of prosodic practices to situate their utterances in relation to the utterances of the others. Through formatting their assessments and response cries into prosodically stylized structures, the interactants co-participate not only in their assessments, but in their "emotional" responses as well. Through a heightened orientation to the prosodic format of the other participants, collaborative highlighting and stylization in assessment sequences allow participants to turn this activity into one of high involvement.

4.2 Prosodic Stylization in Assessments in a Playful Modality: an Independent Gloss of Preference

Auer/Couper-Kuhlen/Müller (1999) argue that interactional rhythm is an independent gloss of preference. Not only rhythm, but prosody in general can be seen in this way. Marked and unmarked prosody – in a similar way as marked and unmarked rhythm – may configure with preferred and dispreferred status in congruent and non-congruent ways. The diagram suggested by Auer/Couper-Kuhlen/Müller for interactional rhythm may also be seen as applicable for marked and unmarked prosody:



(Adapted from Auer/Couper-Kuhlen/Müller, 1999:108)

Marked prosody is defined here as a prosodic gestalt which is heard as decidedly different from the prosodic gestalt in the preceding turn. Marked prosody may engender a stylized prosodic gestalt if the previous turn is nonstylized, but it may also engender a nonstylized gestalt if the previous turn is stylized. In most cases (but certainly not all) marked prosody – like marked interactional rhythm – lines up with dispreferred status; unmarked prosody lines up with preferred status.

That stylised prosody is usually marked is readily apparent. Nonstylized prosody which is marked may be observed in the following example:

4.2.1 “A Whole Other Life”

In an informal conversation, Alina (A) tells her cousin Lenore (L) about a relative, Lisabeth, who has visited Alina’s mother.

1 A: I mean its
2 Daddy’s been dead what
3 seven years
4 and finally it dawns on Lisabeth
5 that she doesn't see mom that much
6 well
7 it’s cause she n
8 I know
9 she never calls her
10 right

For an example of a marked stylized gestalt in an assessment consider the following:

4.2.2 “Confused”

P: is it can we can
it's because you're confused
→ B: <loud, harsh voice, deep fall <no>>
<<audience laughter>>

Following Peter's prosodically non-stylized first turn, Barbara responds with a prosodically stylized “no”, realized with loud volume, harsh voice³³, and with a steep fall into the lower part of her voice range. Similar to beat-clashes, stylized prosody contextualizes emphasis and intensity, and gives Barbara's response added urgency. She is rewarded with audience laughter.

In contrast, the lack of prosodic stylization in a second turn may also be heard as marked. Not reciprocating the prosodic stylization of a first assessment frequently engenders disagreement. Consider the following:

4.2.3 “Gorgeous Girl”

A: Oh: that Pat isn't she a doll::
B: Yeah isn't she pretty
A: Oh she's a gorgeous girl

In this example A creates a stylized gestalt through lengthening on the syllables “oh” and “doll”. B does not reciprocate the prosodic design, which makes it therefore marked. Her unstylized prosodically marked response reinforces the lexical downgrade from the initial assessment. In the third line, A lexically upgrades B's second assessment, prosodically though, it is less stylized than her initial assessment. Lexically, she maintains the same position as in her initial assessment, prosodically, she is reciprocating B's prosodic design in line 2.

³³ For a phonetic description of “harsh voice” see Laver (1980:126-132).

In a formal way, marked prosody – similar to marked interactional rhythm – will engender dispreferred status. By not reciprocating the prosodic design of the first turn assessment, the second speaker has a powerful tool for nonaligning his/her turn with the one preceding. Marked prosody need not always be congruent with dispreferred status, however. Auer/Couper-Kuhlen/Müller (1999) show this with marked rhythm in hot-news informings, in which marked rhythm contextualizes the unexpectedness of the information. It is therefore heard as the preferred response. Within a playful modality, what is usually dispreferred becomes preferred. Marked prosody, then, becomes the preferred response in assessment sequences in which the interactants can be heard to be bantering with one another.

An often overlooked fact in CA research is that it is frequently fun (and preferred) to threaten face, particularly in a play modality. In the right context, threatening face can provide a format for encouraging creative, problem solving second turns by the “threatened participant”. Indeed in one prominent theory of play (Bateson 1954, 1972) a main characteristic of play is aggressive behavior which is signaled and interpreted within the frame “this is play”. What better way to initiate play than by threatening face in an initial assessment and provoking a creative playful response in the second turn?³⁴

As an example of such “face-threatening” behavior in an initial assessment, consider the following example. The excerpt occurs shortly before the above example, “Appease the Monster”. Kendra is about to unwrap her birthday presents. Before she unwraps the first present, she reads the accompanying birthday card, given to her by Wendy and Kevin, who along with Kendra’s parents Marci and Ken, are present:

4.2.4 “Dorky Birthday Card”

```
1 K:  wishing you a day of little treasures
      simple pleasures
      happy birthday
      (.)
→    they're always dorky but at least they have cats on=
      it right
```

³⁴ In her analysis of teasing sequences (*Frotzelnsequenzen*), Günthner (1996, 1998) has described a similar phenomenon.

(.)
 ? : hh.
 → W: <low [Kendra
 [<glass put down>
 10 V M: <<laugh>>
 → W: <high, loud <I'll let you know I spent hours picking=
 =that>>
 [<low, soft <no not really>>
 15 K: [<<smiley voice<I bet you did>>
 M: [<<laugh>>

After reading the poem on the card, Kendra offers a provocative assessment: “they’re always dorky but at least they have cats on it right”. After the negative assessment, she further provokes a response from the other participants when she asks “right”. Following a pause, audible out-breathing, Wendy, simultaneously with the clanking down of a glass, says, “Kendra”, heard as lower than the previous turn. This is followed by laughter by Kevin and Marci and Wendy’s marked stylized response: “I’ll let you know I spent hours picking that”, which is realized with higher volume and higher pitch than the preceding utterances. This prosodically stylized and marked turn is followed by laughter from Marci (line 14) and Kendra’s response, heard with “smiley voice”: “I bet you did” (line 13).

Proof that this sequence takes place in a playful modality is not only Kevin and Marci’s laughter in lines 10 and 14, but also Wendy’s utterance in line 12, “no, not really”, in which she takes back her marked stylized utterance in line 11: “I spent hours picking that”. Her recognition in line 12 of the prior playful modality corresponds with a move back to nonstylized prosody. This is heard as a switch from playful to nonplayful modality.

For further kinds of prosodic stylization within assessment sequences in a playful modality consider the following example:

4.2.5 “Compassionate”

This excerpt is taken from a recording of the radio talk-show program, The Barbara Carlson and Friends Show. Present are Barbara Carlson (B), her co-moderator Peter (P), Senator Graham (G), and a studio audience. Barbara wants to respond to the question how she can be socially liberal and fiscally conservative.

1 B: he wanted to know
I suppose you want to know too
how I can be socially liberal and fiscally=
=conservative

5 well [let me tell you
→ P: [can we say can
→ It's because you're confused
B: no
→ <<audience laughter>>
→ B: it has nothing to do with confusion
it has to do with being compassionate
that's what it [has to do with
G: [a compassionate conservative
P: and Barbara

15 B: yes
P: when you yelled compassionate
it adds that sort of feeling of love that you uh (h)
that you have for all
B: oh peter
10 you are a little turkey

Barbara wants to respond to the question how she can be socially liberal and fiscally conservative. Peter in line 5 begins a floor fight with Barbara, speaking louder: “can we say can”. Barbara relinquishes the floor and Peter offers his assessment: “it’s because you’re confused”. This negative assessment poses a face-threat to Barbara, who, following a pause, answers loudly with a single held syllable, “no”. Her response is followed by audience laughter. Barbara’s rejection of Peter’s assessment is then followed by a prosodic tour de force.

Following the audience laughter Barbara, in a local isochronous rhythm, says:

it has
/nothing to /
/do with con=/
/=fusion /

This is part of the more global isochronous gestalt:

```
it has  
/nothing to do with con= /  
/fusion it /  
/has to do with being com=/  
/passionate /
```

This polyisochronous gestalt³⁵ culminates in Barbara's prosodically upgraded reassessment: "compassionate", which is realized with high volume, high pitch and harsh voice.

A few lines later Peter rewards Barbara's performance, commenting on her prosody:

```
P: when you yelled compassionate  
it adds that sort of feeling of love that you uh (h)  
that you have for all
```

This is seen as proof that the prosodic tour de force by Barbara is not just an analytic concept, but is shown to be relevant for the participants themselves. Barbara's handling of Peter's face-threat within the playful modality shows her masterful management a difficult task.

As stated above, in playful modalities face-threats are a common strategy for encouraging creative responses. Indeed even the casting of the Barbara Carlson and Friends Show seems to encourage the kind of bantering face-threats that are elegantly and creatively solved through practices such as prosodic stylization. Barbara Carlson - a middle-aged, female, liberal Democrat - is pitted against her co-moderator Peter - a young, male, conservative Republican. Assessments are an activity which allows them to playfully attack one another and respond in a creative way. In the above excerpt, such play is rewarded, both by Peter's meta-communicative comment as well as the studio audience's laughter. These examples show that in playful modalities marked prosody cannot be linked congruently with

³⁵ A more detailed account of polyisochronous gestalts can be found in my Master's Thesis, "Musical Performance in Everyday Spoken Discourse: Heightened Involvement Through Collaborative Prosodic Play and Expression of Affect". Universität Konstanz, 1997.

dispreferred status. Marked prosody, like marked interactional rhythm, is seen as an independent gloss of preference.

5 Summary

In second assessments prosodic stylization can be used in order to indicate upgrades and downgrades even when, lexically, the assessments are similar or even the same. Additionally, my corpus shows that prosodically stylized structures do indeed occur in initial and in subsequent assessments. Collaborative prosodic stylization can be a way for interactants to mutually share in the activity. Prosodically marked first and second assessments – which are normally avoided by interactants due to the risk of provoking a dispreferred response – can function to create involvement among the speakers. Furthermore, prosodically stylized second assessments can be heard as cueing a playful modality.

Conclusion

I would like to conclude where I started – with a quotation from Thomas Carlyle:

All passionate language does of itself become musical – with a finer music than the mere accent; the speech of a man even in zealous anger becomes a chant, a song.

(In Chatwin, 1980)

Everyday speech does indeed resemble music, which Thomas Carlyle observed in the 19th century and which I hope to have shown in the preceding dissertation.

Speakers frequently stylize their utterances in order to achieve certain interactional goals. I have attempted to take the first steps in showing how they do this by describing “highlighted” and “stylized” gestalt patterns. To summarize, “highlighted” patterns are those patterns in which a part of a larger pattern is heightened, thereby making it salient. “Stylized” patterns are those patterns in which an entire gestalt is made salient. I have further introduced three processes for highlighting and stylizing prosodic (as well as visual and musical) gestalts: 1) “reduction”, 2) “adjustment”, and 3) “ornamentation”. The process of reduction suppresses local variation in a gestalt pattern; the process of “adjustment” regularizes local variation, and the process of “ornamentation” overlays additional gestalt features on top of an underlying gestalt pattern. I have argued that these three gestalt processes are the basic underlying principles behind gestalt stylization.

In describing how speakers stylize the speech continuum I have applied Max Wertheimer’s gestalt principles of perception. Whereas Wertheimer describes how they are used in visual gestalts and Lerdahl and Jackendoff describe how they are used in musical gestalts, I have described how they are used in prosodic gestalts. Furthermore, I have shown how these three gestalt principles allow intricate grouping patterns which I have termed “compound”, “nested” and “complex”.

In the field of interactional linguistics researchers have only recently begun to consider stylized prosodic gestalts. One consequence of this is a deficit in the terminological apparatus for discussing such phenomena. With the first part of this dissertation I have taken the first steps in correcting this. It is open to further research to refine and expand the results of this work.

Not only has the dissertation attempted to provide a formal description of stylized gestalt patterns, it has also shown how such stylized prosodic gestalts are interactionally relevant and can be exploited to achieve particular interactional goals. One of these interactional goals is the management of turn taking and floor holding in informing sequences. The prosodic stylization and the placement of *oh* in relation to the first turn informing sequence signals who initiates the next discourse action. I've shown that *ohs* which occur as a nested gestalt within the first turn informing are exploited in order to show the speaker's change of state immediately.

In assessment sequences prosodic stylization in second turns may cue upgrading or downgrading even when the utterance is heard as the same on the lexical level. In first turns prosodic stylization may be used to provoke second speakers, which, in a playful modality, often leads to creative, problem solving responses to potentially face-threatening first turns.

Future research could further analyze the role of prosodic stylization in other discourse activities. As a readily deployable practice, I propose that prosodic stylization is exploited in a wide variety of actions and contexts.

Thomas Carlyle's observation that "musical speech" is connected to emotion has been briefly touched on here. This connection, however, must be further explored in future research. Prosodic stylization seen as a conversational practice offers researchers a powerful way of approaching this phenomenon from a CA point of view. Discourse activities in which emotion and prosodic stylization are both present must be analyzed in order to see what role the practice of stylization is playing in the organization of these activities and the connection it has to the display of emotion.

Appendix: Basic Transcription Conventions

Sequential structure

[] overlap
[]
= quick, immediate connection of new turns or single units

Pauses

(.) micro-pause
(-), (--), (---) short, middle or long pauses of approx. 0.25 - 0.75 seconds, up to approx. 1 second
(2.0) estimated pause of more than approx. 1 second
(2.85) measured pause (measured to hundredths of a second)

Rhythm

/ / / isochronous beats

Other segmental conventions

and=uh slurring within units
:, :;, :: lengthening, according to its duration
uh, ah, etc. hesitation signals, so-called "filled pauses"
' glottal stop

Laughter

so(h)o laughing particles during speech
haha hehe hihi syllabic laughing
((laughing)) description of laughter

Reception signals

hm, yes, yeah, no one syllable signals
hm=hm, yea=ah, no=o two syllable signals
'hm'hm two syllable signal with a glottal stop, usually signals negation

Accents

ACcent primary, or main accent
!AC!cent extra strong accent

Final pitch movements

? high rise
, mid-rise
- level pitch
; mid-fall
. low fall

Pitch step-up/step down

↓ pitch step down
↑ pitch step up

Change of pitch register

<<l> > low pitch register
<<h> > high pitch register

Volume and tempo changes

<<f> > forte, loud
<<ff> > fortissimo, very loud
<<p> > piano, soft
<<pp> > pianissimo, very soft
<<all> > allegro, fast
<<ten> > lento, slow
<<cresc> > crescendo, becoming louder
<<dim> > diminuendo, becoming softer
<<acc> > accelerando, becoming faster
<<rall> > rallentando, becoming slower

Breathing in and out

.h, .hh, .hhh breathing in, according to its duration
h, hh, hhh breathing out, according to its duration

Other conventions

((cough)) paralinguistic and non-linguistic actions and events
<<coughing> > accompanying paralinguistic and non-linguistic actions and events over a stretch of speech
<<surprised> > interpretive comments over a stretch of speech
() unintelligible passage, according to its duration
(such) presumed wording
al(s)o presumed sound or syllable
(such/which) possible alternatives
(()) omission of text
→ specific line in the transcript which is referred to in the text

Bibliography

- Ashby, M. G. (1978) "A Study of Two English Nuclear Tones". In: Language and Speech. 21: pp. 326-336.
- Atkinson, J.M./Heritage, John, eds. (1984) Structures of Social Action: Studies in Conversational Analysis. Cambridge: Cambridge University Press.
- Auer, Peter (1986) "Kontextualisierung". In: Studium Linguistik 19: pp. 22-47.
- Auer, Peter (1989) "Natürlichkeit und Stil". In: Volker Hinnenkamp, Margret Selting, (eds.): Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik. Tübingen: Max Niemeyer Verlag, pp. 27-59.
- Auer, Peter (1996) "On the Prosody and Syntax of Turn-Continuations". In: Elizabeth Couper-Kuhlen/Margret Selting, (eds.): Prosody in Conversation: Interactional Studies. Cambridge: Cambridge University Press: pp. 57-100.
- Auer, Peter/Couper-Kuhlen, Elizabeth (1994) "Rhythmus und Tempo konversationeller Alltagssprache". In: Zeitschrift für Literaturwissenschaft und Linguistik. 96: pp. 78-106.
- Auer, Peter/Couper-Kuhlen, Elizabeth/Müller, Frank (1999) Language in Time: The Rhythm and Tempo of Spoken Interaction. New York, Oxford: Oxford University Press.
- Auer, Peter/di Luzio, Aldo (eds). (1992) The Contextualization of Language. Amsterdam/Philadelphia: Benjamins.
- Auer, Peter/Uhmann, Susanne (1982). „Aspekte der konversationellen Organisation von Bewertungen“. In: Hugo Steger/Odo Leys/Gerhard Stichel/Johannes Schwitalla (eds.): Deutsche Sprache: Zeitschrift für Theorie, Praxis, Dokumentation. pp. 1-32.
- Bateson, Gregory (1954) A Theory of Play and Fantasy. Steps to an Ecology of Mind. New York: Ballentine.
- Bateson, Gregory (1972) Steps to an Ecology of Mind. New York: Ballentine Books.
- Bolinger, Dwight L. (1958) "Intonation and Grammar". In: Language Learning. 8: pp. 31-37.
- Brazil, D./Coulthard, M./Johns, C. (1980) Discourse Intonation and Language Teaching. London: Longman.

- Brown, G./Currie, K./Kenworthy, J. (1980) Questions of Intonation. London: Croom Helm.
- Chatwin, Bruce (1980) The Songlines. London: Jonathon Cape, Ltd.
- Cruttenden, A. (1986) Intonation. Cambridge: Cambridge University Press.
- Cook-Gumperz, Jenny/Gumperz, John J. (1976) "Context in Children's Speech". In: Papers on Language and Context. Working Paper 46. Berkley: Language Behavior Research Laboratory.
- Cooper, Pete (1995) Mel Bay's Complete Irish Fiddle Player. Pacific, MO: Mel Bay Publications.
- Couper-Kuhlen, Elizabeth (1986) An Introduction to English Prosody. Tübingen: Niemeyer.
- Couper-Kuhlen, Elizabeth (1990) „Discovering Rhythm in Conversational English: Perceptual and Acoustic Approaches to the Analysis of Isochrony." In: KontRI: Kontextualisierung durch Rhythmus und Intonation. Nr. 13. Universität Konstanz, September.
- Couper-Kuhlen, Elizabeth (1993) English Speech Rhythm. Form and Function in Everyday Verbal Interaction. Amsterdam: Benjamins.
- Couper-Kuhlen, E./Selting, Margret (1996) „Towards an Interactional Perspective on Prosody and a Prosodic Perspective on Interaction". In: Elizabeth Couper-Kuhlen/Margret Selting (eds.): Prosody in Conversation. Cambridge: Cambridge University Press, pp. 11-56.
- Crystal, David (1969) Prosodic Systems and Intonation in English. Cambridge: Cambridge University Press.
- Crystal, David (1975) The English Tone of Voice: Essays in Intonation, Prosody and Paralanguage. London: Edward Arnold.
- Deppermann, Arnulf (1999) Gespräche analysieren: Eine Einführung in konversationsanalytischen Methoden. Opladen: Leske + Budrich.
- Crystal, D./Davy, D. (1975) Advanced Conversational English. London: Longman.
- DuBois, J. W./Schuetze-Coburn, S./Cumming, S./Paolino, D. (1993) "Outline of Discourse Transcription". In: J. A. Edwards/M. D. Lampert (eds.): Talking Data: Transcription and Coding in Discourse Research. Hillsdale, New Jersey: Lawrence Erlbaum, pp. 45-89.
- Erickson, Frederick (1981) "Money Tree, Lasagna Bush, Salt and Pepper: Social Construction of Topical Cohesion in a Conversation among Italian Americans".

- In: Deborah Tannen (ed.): Analyzing Discourse: Text and Talk. Washington, D.C.: Georgetown University Press: pp. 43-70.
- Erickson, Frederick (1992) "‘They Know All the Lines’: Rhythmic Organization and Contextualization in a Conversational Listing Routine". In: Peter Auer, Aldo di Luzio (eds.): The Contextualization of Language. Amsterdam: Benjamins: pp. 365-398.
- Flowe, William (1997) Musical Performance in Everyday Spoken Discourse: Heightened Involvement through Collaborative Prosodic Play and Expression of Affect. Schriftliche Arbeit für die Anerkennung des "Master of Arts" Diplom als Zulassungsvoraussetzung für eine Promotion an der Universität Konstanz.
- Freese, Jeremy/Maynard, Douglas W. (1998) "Prosodic Features of Bad News and Good News in Conversation". In: Language in Society. 27, pp. 195-219.
- French, P. & Local, J. (1983), "Turn-competitive Incomings", In: Journal of Pragmatics 7, pp. 17-38.
- Garfinkel, Harold (1967) Studies in Ethnomethodology. Englewood Cliffs: Prentice Hall.
- Gavioli, Laura (1995) "Turn-Initiation vs. Turn-Final Laughter: Two Techniques for Initiating Remedy in English/Italian Bookshop Service Encounters". In: Discourse Processes. 19, 3: May-June: pp. 369-384.
- Glenn, Phillip. (1989) "Initiating Shared Laughter in Multi-Party Conversations". In: Western Journal of Speech Communication. 53, 2, Spring: pp. 127-149.
- Glenn, Phillip. (1991-1992) "Current Speaker Initiation of Two-Party Shared Laughter". In: Research on Language & Social Interaction. 25: pp. 139-162.
- Goffman, Erving (1964) "On Face Work: An Analysis of Ritual Elements in Social Interaction". In: W.G. Bennis (ed.): Interpersonal Dynamics. Hoewood, Ill.: The Dorsey Press, pp. 226-249.
- Goffman, Erving (1967) Interactional Ritual. Garden City: Doubleday.
- Goffman, Erving (1974) Frame Analysis. New York: Harper and Row.
- Goffman, Erving (1978) "Response Cries". In: William Bright (ed.): Language 54, 4: pp. 787-815.
- Goffman, Erving (1983) "The Interaction Order". In: American Sociological Review 48: pp. 1-17.

- Goodwin, Charles/Goodwin, Marjorie Harness (1987) "Concurrent Operations on Talk: Notes on the Interactive Organization of Assessments". IPRA Papers in Pragmatics 1, No. 1: pp. 1-54.
- Goodwin, Charles/Goodwin, Marjorie Harness (1992) "Context, Activity and Participation". In: Peter Auer/Aldo di Luzio, (eds.): The Contextualization of Language. Amsterdam: John Benjamins: pp. 77-99.
- Grelling, Kurt/Oppenheim, Paul (1937/38) "Der Gestaltbegriff im Lichte der neuen Logik". Erkenntnis, 7: pp. 211-225. English translation "The concept of gestalt in the light of modern logic". In: Barry Smith (ed.): Foundations of Gestalt Psychology. München, Wien: Philosophia Verlag GmbH, 1988.
- Günthner, Susanne (1996) „Zwischen Scherz und Schmerz – Frotzelaktivitäten in Alltagsinteraktionen.“ In: Helga Kotthof (ed.): Scherzkommunikation. Opladen: Westdeutscher Verlag: pp. 81-108.
- Günthner, Susanne (2000) Vorwürfsaktivitäten in der Alltagsinteraktion: Grammatische, prosodische, rhetorisch-stilistische und interaktive Verfahren bei der Konstitution kommunikativer Muster und Gattungen. Tübingen: Niemeyer.
- Gumperz, John J. (1982) Discourse Strategies. Cambridge, Cambridge University Press.
- Gumperz, John J. (1992) "Contextualization and Understanding". In: Duranti, Alessandro/Goodwin, Charles (eds.) Rethinking Context. Language as an Interactive Phenomenon. Cambridge/New York: Cambridge University Press: pp. 229-252.
- Gumperz, John J. (1992a) "Contextualization Revisited". In: Auer, Peter/di Luzio, Aldo (eds.) The Contextualization of Language. Amsterdam/Philadelphia: Benjamins: pp. 39-45.
- Gumperz, John J. (1996) „Forward by John Gumperz“. In: Elizabeth Couper-Kuhlen/Margret Selting (eds.) Prosody in Conversation. Cambridge: Cambridge University Press: pp. x-xii.
- Gumperz, John J./Hymes, Dell (1972) "Remarks on Ehnomethodology". In: John J. Gumperz/Dell Hymes (eds.) Directions in Sociolinguistics. New York: Holt, Rinehart & Winston.
- Gumperz, John J./Levinson, Stephen C. (1996) Rethinking Linguistic Relativity. Cambridge: Cambridge University Press.

- Gumbrecht, H. U. (1986) „Schwindende Stabilität der Wirklichkeit. Eine Geschichte des Stilbegriffs“. In: H.U. Gumbrecht/Armin Biermann (eds.): Stil. Geschichten und Funktionen eines kulturwissenschaftlichen Diskurselements. Frankfurt: Suhrkamp: pp. 726-788.
- Haiman, John (1994) "Ritualization and the Development of Language." In: William Pagliuca (ed.): Perspectives on Grammaticalization. Amsterdam: Benjamins: pp. 5-28.
- Halliday, M. A. K. (1967) Intonation and Grammar in British English. The Hague, NL: Mouton.
- Halliday, M. A. K. (1970) A Course in Spoken English: Intonation. London: Oxford University Press.
- Heritage, John (1984) "A Change of State Token and Aspects of its Sequential Placement." In: J. M. Atkinson/J. Heritage (eds.) Structures of Social Action: Studies in Conversation Analysis. Cambridge: Cambridge University Press: pp. 299-345.
- Heritage, John. (1998) "Oh-Prefaced Responses to Inquiry". In: Language in Society. 27: pp. 291-334.
- Hickey, Leo (1989) The Pragmatics of Style. London, New York: Routledge.
- Hinnenkamp, Volker (1989) "Die Stilisierung von Ethnizität". In: Volker Hinnenkamp/Margret Selting (eds.): Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik. Tübingen: Max Niemeyer Verlag: pp. 253-291.
- Hymes, Dell (1974) "Ways of Speaking". In: Richard Baumann/Joel Scherzer (eds.) Explorations in the Ethnography of Speaking. New York, London: Cambridge University Press. Translation: "Über Sprechweisen". In: Soziolinguistik Zur Ethnographie der Kommunikation. Florian Coulmas (ed.). Frankfurt: Suhrkamp, 1979: pp. 166-192.
- Jefferson, Gail (1978) "Sequential Aspects of Storytelling in Conversation" In: Jim Schenkein (ed.) Studies in the Organization of Conversational Interaction, New York: Academic Press: pp. 219-248.
- Jefferson, Gail. (1979): "A Technique for Inviting Laughter and its Subsequent Acceptance Declination". In: George Psathas (ed.): Everyday Language. Studies in Ethnomethodology. New York: Irvington Publishers; pp. 79-96.
- Jefferson, Gail (1984): "On the Organization of Laughter in Talk about Troubles". In: J. Maxwell Atkinson/John Heritage (eds). (1984): Structures of Social

- Interaction: Studies in Conversation Analysis. Cambridge: Cambridge University Press: pp. 346-369.
- Keim, Imken/Schwittala, Johannes (1989) „Soziale Stile des Miteinandersprechens. Beobachtungen zu Formen der Konfliktbearbeitung in zwei Frauengruppen“. In: Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik. Volker Hinnenkamp/Margret Selting, (eds.). Tübingen: Max Niemeyer Verlag: pp. 83-121.
- Klewitz, Gabriele/Couper-Kuhlen Elizabeth (1999) „Quote-Unquote? The Role of Prosody in the Contextualization of Reported Speech Sequences“. Pragmatics 9.4: 459-485.
- Koffka, K. (1922) „Perception: An Introduction to the Gestalttheorie“, Psychological Bulletin. 19: pp. 531-85. Reprinted in T. Shipley (ed.) Classics in Psychology. New York: Philosophical Library, 1961: pp. 1128-1196.
- Koffka, K. (1923) “Über die Untersuchungen an den sogenannten optischen Anschauungsbildern. Forschungsbericht“. In: Psychologische Forschung. 3: pp. 124-167.
- Koffka, K. (1924) „New Experiments in the Perception of Movement“ In: C.S. Myers (ed.) Proceedings and Papers of the 7th International Congress of Psychology at Oxford 1923. Cambridge, Cambridge University Press: pp. 369-373.
- Koffka, K. (1924/1925) “Psychical and Physical Structures“. In: Psyche. 5: pp. 80-85.
- Koffka, K. (1928) “On Gestalt Theory“. In: The Smith Alumnae Quarterly. Feb. 1928: pp. 142-147.
- Koffka, K. (1931) “Die Wahrnehmung von Bewegung“. In: A. Bethe, (ed.) Handbuch der normalen und pathologischen Physiologie. 12/2, Berlin: Springer: pp. 1166-1214.
- Koffka, K. (1935) Principles of Gestalt Psychology. New York: Harcourt, Brace, and Comp., London: Kegan Paul, repr. 1963.
- Köhler, W. (1920) Die physischen Gestalten in Ruhe und im stationären Zustand. Eine naturphilosophische Untersuchung. Vieweg, Braunschweig. Translated extract „Physical Gestalten“ in W. D. Ellis (ed.) A Source Book of Gestalt Psychology. London: Routledge & Kegan Paul: pp. 17-54.
- Köhler, W. (1923) “Zur Theorie der stroboskopischen Bewegung (Im Anschluß an die Untersuchung Hartmanns)“. In: Psychologische Forschung. 4: pp. 115-175.

- Köhler, W. (1924) "The Problem of Form in Perception". In: C.S. Myers (ed.) Proceedings of the 7th International Congress of Psychology. Cambridge: Cambridge University Press: pp. 22-28.
- Köhler, W. (1925) "An Aspect of Gestalt Psychology". In C. Murchison (ed.) Psychologies of 1925. Worcester, Mass.: Clark University Press.
- Köhler, W. (1925a) "Komplextheorie und Gestalttheorie. Antwort auf G. E. Müllers Schrift gleichen Namens". In: Psychologische Forschung. 6: pp. 258-416. Translated extract „Reply to G.E. Müller". In W.D. Ellis (ed.) A Source Book of Gestalt Psychology. London: Routledge & Kegan Paul, 1938: pp. 379-388.
- Köhler, W. (1929) Gestalt Psychology. New York: Liveright,
- Köhler, W. (1929a) "Ein altes Scheinproblem". In Die Naturwissenschaften. 17: pp. 395-401.
- Köhler, W. (1930) "Some Tasks of Gestalt Psychology". In C. Murchison (ed.) Psychologies of 1930. Worcester, Mass.: Clark University Press: pp. 143-160.
- Kotthof, Helga (1989) „Stilunterschiede in argumentativen Gesprächen oder zum Geselligkeitwert von Dissens“. In: Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik. Volker Hinnenkamp/Margret Selting (eds.) Tübingen: Max Niemeyer Verlag: pp. 187-202.
- Ladd, D. Robert/Scherer K. R./Silverman, Kim (1986) "An Integrated Approach to Studying Intonation and Attitude". In: Catherine Johns-Lewis (ed.) Intonation in Discourse. London & Sydney: Croom Helm.
- Laver, John (1980) The Phonetic Description of Voice Quality. Cambridge: Cambridge University Press.
- Laver, John (1994) Principles of Phonetics. Cambridge: Cambridge University Press.
- Leech G/Svartvik, J. (1975) A Communicative Grammar of English. London: Longman.
- Lerdahl, F.; Jackendoff, R. (1983) A Generative Theory of Tonal Music. Cambridge, Mass.: MIT Press.
- Local, John (1986) "Patterns and problems in a study of Tyneside intonation". In C. Johns-lewis (ed.) Intonation and Discourse. London: Croom Helm: pp. 181-198.
- Local, John (1996) "Conversational Phonetics: Some Aspects of News Receipts in Everyday Talk". In Margret Selting/Elizabeth Couper-Kuhlen (eds.) Prosody in Conversation: Interactional Studies. Cambridge: Cambridge University Press: pp. 177-230.

- Local, J.K., J. Kelly and W. H. G. Wells (1986) "Towards a phonology of conversation: turn taking in Tyneside English". In: Journal of Linguistics. 22:411-437.
- Mach, Ernst (1866) Beiträge zur Analyse der Empfindungen. 1st edition. Jena: Fischer; English translation: Contributions to the Analysis of the Sensations by C. M. Williams, Chicago: Open Court, 1897.
- Metzger, Wolfgang. (1954) „Grundbegriffe der Psychologie“. In: Michael Stadler/Heinrich Crabus (eds.) Gestalt-Psychologie: Ausgewählte Werke aus den Jahren 1950 bis 1982. Frankfurt am Main: Waldemar Kramer Verlag. 1986: pp. 124-133.
- Müller, Frank (1989) „Isometrische Strukturen des Italienischen.“ In: KontRI: Kontextualisierung durch Rhythmus und Intonation. Nr. 10. Universität Konstanz, September.
- Müller, Frank (1989a) „Lautstilistische Muster in Alltagstexten von Südtalienern“ In: Volker Hinnenkamp/Margret Selting (eds.) Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik. Tübingen: Max Niemeyer Verlag: pp. 61-81.
- Müller, Frank (1991) „Metrical Emphasis: Rhythmic Scansions in Italian Conversation“. KontRI: Kontextualisierung durch Rhythmus und Intonation. Nr. 14. Universität Konstanz, February.
- Munby, J. (1978) Communicative Syllabus Design, Cambridge, Cambridge University Press.
- Ono, Tsuyoshi/Thompson, Sandra A. (1995) "What Can Conversation Tell us about Syntax?" In: P.W. Davis (ed.) Alternative Linguistics. Amsterdam, Benjamins: pp. 213-271.
- Panese, Marcello (1992) "Calling In: Sequential and Prosodic Aspects of Openings in Radio Talk." KontRI: Kontextualisierung durch Rhythmus und Intonation. Nr. 20. Universität Konstanz, Fachgruppe Sprachwissenschaft.
- Pomerantz, Anita (1975) Second Assessments. A Study of Some Features of Agreements/Disagreements. Irvine: University of California at Irvine.
- Pomerantz, Anita (1984) "Agreeing and Disagreeing with Assessments: Some Features of Preferred/Dispreferred Turn Shapes". In: J. M Atkinson/J. Heritage (eds.) Structures of Social Action: Studies in Conversation Analysis. Cambridge: Cambridge University Press: pp. 57-101.

- Paul, Ingwer (1989) "Ritual und Ritualszitat". In: Volker Hinnenkamp/Margret Selting (eds.) Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik. Tübingen: Max Niemeyer Verlag: pp. 167-185.
- Quirk, R./Greenbaum, S./Leech, G./Svartvik, J. (1972) A Grammar of Contemporary English. London: Longman.
- Sacks, H. (1992) Lectures on Conversation. Oxford: Blackwell.
- Sacks, H./Schegloff, E. (1973) „Opening up Closings“. In: Semiotica. 8: pp. 289-327.
- Sacks, H./Schegloff, E./Jefferson, G. (1974) "A Simplest Systematics for the Organization of Turn-Taking for Conversation". In: Language. 50: pp. 696-735.
- Sacks, H./Schegloff, E./Jefferson, G. (1977) "The Preference for Self-Correction in the Organization of Repair in Conversation". Language 53, pp. 361-382.
- Schegloff, Emanuel, A. (1983) "Reflections on Quantification in the Study of Conversation". In: Research on Language & Social Interaction. 26, 1: pp. 99-128.
- Schegloff, Emanuel, A. (1996) "Confirming Allusions: Towards an Empirical Account of Action". In: American Journal of Sociology. 102, 1, July: pp. 161-216.
- Schegloff, Emanuel, A. (1997) "Practices and Actions: Boundary Cases of Other-Initiated Repair". In: Discourse Processes. 23: pp. 499-547.
- Schegloff, Emanuel, A. (1998) "Reflections on Studying Prosody in Talk-in-Interaction." In: Language and Speech. 41 (3-4): pp. 235-263.
- Schütz, Alfred (1932) Der sinnhafte Aufbau der sozialen Welt. Wien. Reprint (1984) Frankfurt am Main: Suhrkamp.
- Schütz, Alfred/Luckmann, Thomas (1975) Strukturen der Lebenswelt. Frankfurt am Main: Suhrkamp.
- Schuetze-Coburn, Stephan/Shapley, Marian/Weber,Elizabeth (1991) "Units of Intonation in Discourse: A Comparison of Acoustic and Auditory Analyses" In: Language and Speech. 34, 3: pp. 207-243.
- Selting, Margret (1989) "Konstitution und Veränderung von Sprechstilen als Kontextualisierungsverfahren: Die Rolle von Sprachvariation und Prosodie". In: Volker Hinnenkamp/Margret Selting (eds.) Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik. Tübingen: Max Niemeyer Verlag: pp. 203-225.
- Selting, Margret (1992a) „Prosody in Conversational Questions“. In: Journal of Pragmatics. 17: 315-345.

- Selting, Margret (1992b) „Intonation as a Contextualization Device: Case Studies on the Role of Prosody, Especially Intonation in Contextualizing Storytelling in Conversation“. In: Peter Auer/Aldo di Luzio (eds.) The Contextualization of Language. Amsterdam and Philadelphia: Benjamins: pp. 233-258.
- Selting, Margret (1994) „Emphatic Speech Style – With a Focus on the Prosodic Signalling of Heightened Emotive Involvement in Conversation“. In: Journal of Pragmatics. 22: pp. 375-408.
- Selting, Margret (1995) Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation. Tübingen: Niemayer.
- Selting, Margaret (1996) „On the Interplay of Syntax and Prosody in the Constitution of Turn Constructional Units and Turns in Conversation“. In: Pragmatics. 6: pp. 357-388.
- Selting, Margret/Sandig, Barbara (eds.) (1997) Sprech- und Gesprächsstile. Berlin, New York: De Gruyter.
- Selting, Margret/Auer, Peter/Barden Birgit/Couper-Kuhlen, Elizabeth/Günthner, Susanne/Quasthoff, Uta M./Schlobinski, Peter/Uhmann, Susanne (1998) „Gesprächsanalytisches Transkriptionssystem (GAT)“. In: Linguistische Berichte. 34: pp. 91-122.
- Szczepek, Beatrice (1998) On the Prosody and Syntax of Collaborative Productions in Spoken English. Unpublished Masters Thesis. Universität Konstanz.
- Szczepek, Beatrice (2000) “Formal Aspects of Collaborative Productions in English Conversation“. In: InList: Interaction and Linguistic Structures. 17. Konstanz: Universität Konstanz, June.
- Tench, P. (1990) The Roles of Intonation in English Discourse. Frankfurt am Main: Peter Lang.
- Uhmann, Susanne (1989) “Interviewstil: Konversationelle Eigenschaften eines sozialwissenschaftlichen Erhebungsinstruments“. In: Volker Hinnenkamp/Margret Selting (eds.) Stil und Stilisierung: Arbeiten zur interpretativen Soziolinguistik. Tübingen: Max Niemeyer Verlag: pp. 125-165.
- Uhmann, Susanne (1996) “On Rhythm in Everyday German Conversation: Beatclashes in Assessment Utterances“. In: Prosody in Conversation. Elizabeth Couper-Kuhlen, Margret Selting, eds. Cambridge University Press. pp. 303-365.
- Von Ehrenfels, Christian (1890) „Über Gestaltqualitäten“ In: Vierteljahrsschrift für wissenschaftliche Philosophie. 14: pp. 242-292. English Translation: „On

- ‘Gestalt Qualities’“. In: Barry Smith (ed.) Foundations of Gestalt Theory. München, Wien: Philosophia Verlag, 1988: pp. 82-117,.
- Von Ehrenfels, Christian (1916) Kosmogonie. Jena: Diederichs: pp. 93-96. English Translation „Gestalt Level and Gestalt Purity“. In: Barry Smith (ed.) Foundations of Gestalt Psychology. München, Wien: Philosophia Verlag, 1988: pp. 118-120.
- Van Ek, J. (1975) Systems Development in Adult Language Learning: The Threshold Level. Strasbourg: Council of Europe.
- Wertheimer, M. (1922) „Untersuchungen zur Lehre von der Gestalt I. Prinzipielle Bemerkungen“. In: Psychologische Forschung. 12: pp. 47-58.
- Wertheimer, M. (1923) “Untersuchungen zur Lehre von der Gestalt II”. In: Psychologische Forschung. 4: pp. 301-350. English Translation: “Laws of Organization in Perceptual Forms”. In: Willis D. Ellis (ed.) A Source Book of Gestalt Psychology. London: Routledge & Kegan Paul, 1938: pp. 71-88.
- Wertheimer, M. (1925) „Über Gestalttheorie“. Erlangen: Philosophische Akademie, English Translation: „Gestalt Theory“ In: Willis D. Ellis (ed.) A Source Book of Gestalt Psychology. London: Routledge & Kegan Paul, 1938: pp. 1-11.
- Wertheimer, M. (1925a) Drei Abhandlungen zur Gestalttheorie. Erlangen: Philosophische Akademie.
- Wertheimer, M. (1927) „Gestaltpsychologische Forschung“. In: E. Saube (ed.) Einführung in die neuere Psychologie: Handbuch der neueren Erziehungswissenschaften. Osterwieck/Harz: Zwickfeldt: pp. 46-53.
- Wertheimer, M. (1933) „Zu dem Problem der Unterscheidung von Einzelinhalt und Teil“. In: Zeitschrift für Psychologie. 129: pp. 353-357.
- Wertheimer, M. (1938) „Gestalt Theory“. In: Willis D. Ellis (ed.) A Source Book of Gestalt Psychology. London: Routledge & Kegan Paul: pp. 1-11.
- Wilkins, D.A. (1976) Notional Syllabuses. London, Oxford University Press.