

Kontrastprosodie in Sprachproduktion und -perzeption

Kai Alter*, Ina Mleinek, Tobias Rohe*, Anita Steube, Carla Umbach
Universität Leipzig

*Max-Planck-Institut für neuropsychologische Forschung

1 Einleitung

Die vorliegende Studie zum Deutschen verfolgt zwei Ziele zur Untersuchung des Zusammenwirkens von Semantik und Prosodie:

- (1) Wie wird Kontrast in der spezifischen Form von Korrektur prosodisch/akustisch realisiert, und
- (2) wie nutzt der Hörer spezifische prosodische/akustische Merkmale der Kontrastprosodie für deren Interpretation.

Eine Reihe von Untersuchungen im Rahmen der Sprachproduktion von Kontrast/Korrektur konnten im Deutschen wie auch in anderen Sprachen belegen, dass die Realisierung von Kontrast unmittelbar an eine spezifische Ausbuchstabierung von akustischen Parametern gebunden ist (Alter et al., 1997; Bolinger, 1961; Cooper et al., 1985; Garding, 1981; Dietrich, 1990; Mehlhorn & Zybatow, 2000; Mehlhorn, 2001; Zybatow & Mehlhorn, 2000).

Die hier vorgestellten Studien zielen auf die Produktion **und** Perzeption von Kontrast/Korrektur ab. Die prosodischen Daten ermittelten wir anhand von signalakustischen Analysen gesprochener Sprache. Ein Novum an unserer Studie ist, dass wir die prosodische Realisierung von Kontrast/Korrektur auch in einem auditorischen Verhaltenstest verwendeten, um somit zusätzliche Erkenntnis über die Gewichtigkeit prosodischer Parameter während der Perzeption zu gewinnen.

Der legendäre Frage-Antwort-Test ist durchaus geeignet, Hintergrund-Fokus-Aufteilungen von Textsätzen zu erzeugen oder Kontrastakzente zu erzwingen. Dieser Test muss aber Restriktionen unterliegen, wenn es darum geht, die semantische Unterspezifikation ganzer Textteile hinsichtlich einer möglichen Korrekturinterpretation zu ermitteln (vgl. 4.3, s. auch Steube im Druck; Steube in diesem Band). Deshalb haben wir uns entschieden, mittels kleiner Anekdoten und der damit einhergehenden Kontextinformation zwei sowohl semantisch als auch prosodisch vergleichbare Sprecher- als auch Hörersituationen zu schaffen, die einerseits auf weiten Neuinformationsfokus und andererseits auf (Korrektur)-Kontrastfokus abzielen. Damit haben wir sowohl hinsichtlich der Sprachproduktion als auch der Sprachperzeption

eine schwierigere Situation für die Probanden geschaffen (Allbritton et al., 1996), als sie durch die gängigen Frage-Antwort-Sequenzen erzeugt werden würde.

2 Design der Produktionsstudie

2.1 Kontexte und Targetsätze

Insgesamt wurden 16 verschiedene Kontexte konstruiert, von denen 8 Kontexte Neuinformationsfokus (NIF) und 8 Kontexte Kontrastfokus (KF) induzieren. Die Besonderheit an unserer Studie ist, dass die Position von Neuinformationsfokus mit derjenigen von Kontrastfokus übereinstimmt. Aus prosodischer Sicht führt das dazu, dass die herkömmliche Satzposition mit der Kontrastposition lokal zusammenfällt. Unter kanonischer Akzentposition verstehen wir diejenige Akzentposition, wie sie im Fall von NIF auf Satzebene realisiert wird.

Unsere Hypothese ist, daß ein Satz - eingebettet in unterschiedliche Kontexte, aber mit gleicher linearer Wortfolge und gleicher syntaktischer Struktur – (unabhängig von individuellen Sprecherrealisierungen) konstant divergierende prosodische Realisierungen erfährt. Unterschiedliche Kontextinformationen sollten auch zu unterschiedlichen prosodischen Mustern führen.

Im Folgenden präsentieren wir ein Beispiel für solch ein kontextuelles Minimalpaar. Beginnen wir zunächst mit Beispiel (1)(a) für 'Neuinformationsfokus':

- (1) (a) Semesterende... Martina ist durch die Prüfung gefallen und traut sich nicht, nach Hause zu ihren Eltern zu fahren. Andererseits sind die meisten aus ihrem Wohnheim inzwischen abgereist und die Flure sind deprimierend leer. Endlich fasste Martina einen Entschluss ...
(Hörprobe):
Am Samstag Morgen packte Martina ihre Reisetasche. Sie fuhr zum Bahnhof. Dort kaufte sie eine Fahrkarte nach Paris und setzte sich in die Wartehalle...

Von dem in dieser kleinen Anekdote unterstrichen dargestellten Satz erwarten wir, dass er den Kriterien eines Neuinformationsfokus genügt. Baut man nun um diesen Satz einen anderen Kontext, nämlich einen Kontext, der Kontrast auslöst, wäre zu erwarten, dass bei gleicher Wortfolge dieses Satzes eine Änderung in der prosodischen Realisierung auftritt.

Betrachten wir Beispiel (1)(b) mit 'Kontrastfokus':

- (1) (b) Martinas Eltern machen sich Sorgen, weil Martina nicht nach Hause gekommen ist. Nach viel Mühe schaffen sie es, eine Studentin aus dem Wohnheim an das Telefon zu bekommen, die Martina vom Sehen her kennt. Die Verbindung ist schlecht.

Studentin am Telefon:

(Hörprobe):

Martina? Die habe ich doch am Samstag noch im Bus gesehen! Wo sie hin gefahren ist? ... Nein, nicht in die Innenstadt. Sie fuhr zum Bahnhof.

Wir gehen hier der Frage nach, ob der Zielsatz '*Sie fuhr zum Bahnhof.*' - eingebettet in verschiedene Kontexte - eine unterschiedliche prosodische Realisierung erfährt. Die Positionierung der Hervorhebung/Prominenz geschieht dabei auf jeweils derselben Konstituente, nämlich ‚Bahnhof‘. Sowohl der Satzakzent in der NIF-Bedingung (1)(a) als auch der Kontrastakzent in der KF-Bedingung (1)(b) sind mit derselben lexikalisch betonten Silbe ‚BAHN-‘ des Wortes ‚Bahnhof‘ assoziiert.

2.2 Experimentelles Design

Mittels der 2x8 kleinen Anekdoten wurde eine Sprachproduktionsstudie durchgeführt. Das gesamte Korpus der Sprachproduktionsstudie umfasst 8 Kontexte pro Bedingung, produziert von 8 deutschen Muttersprachlerinnen, und resultiert somit in 64 Targetsätzen pro Bedingung. Die Kontexte einschließlich der Zielsätze wurden mit einem DAT-Recorder aufgenommen und digitalisiert. Die Zielsätze wurden mit Hilfe von Cool aus den Kontexten ausgeschnitten und mittels WinPitch und PitchWorks akustisch und prosodisch analysiert.

3 Akustisch/prosodische Auswertung

Die Fragestellung war, ob es einen Unterschied zwischen der Akzentrealisierung in der Neuinformationsfokusbedingung und derjenigen in der Kontrastfokusbedingung gibt. Betrachten wir zunächst Einzelfälle von akustischen Analysen der Zielsätze, ist zumindest eine erste Bestätigung unserer Erwartung an eine unterschiedliche prosodische Realisierung zu vermerken, vgl. Abbildung 1:

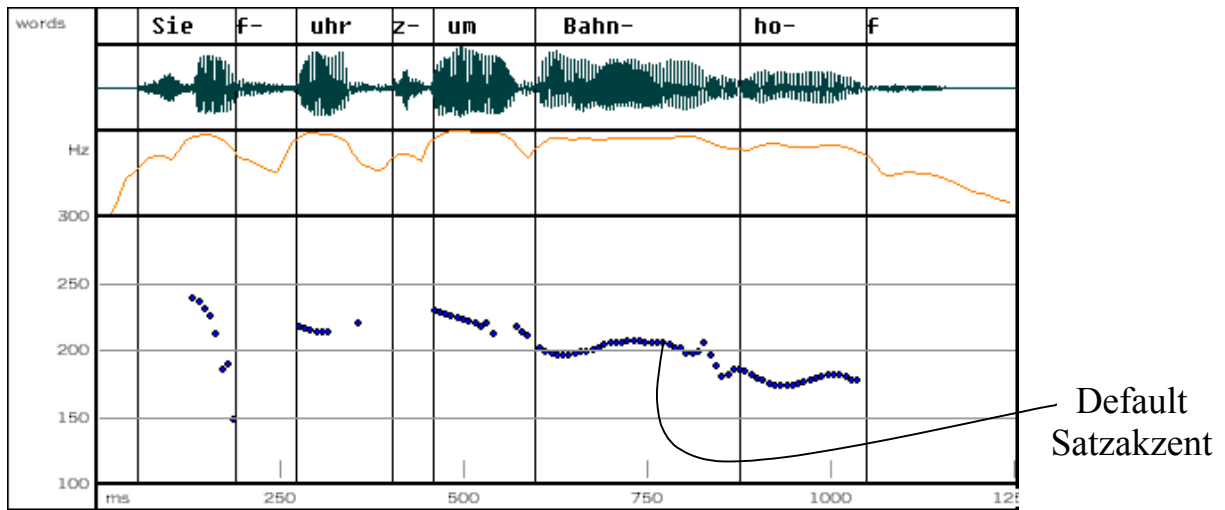


Abb. 1: Ein akustischer Einzelfall des Satzes ‚*Sie fuhr zum Bahnhof.*‘ im Kontext mit Neuinformationsfokus (1)(a).

Zu sehen ist hier eine dreigeteilte Darstellung des akustischen Signals des Satzes ‚*Sie fuhr zum Bahnhof.*‘, realisiert von einer Sprecherin. Im oberen Teil ist der zeitliche Zusammenhang von Satzkonstituenten (Einzelwörter) zum akustischen Signal dargestellt. Unmittelbar darunter folgt das Oszillogramm, mit anschließender Darstellung der Amplitudenhülle. Den Abschluss (unterer Teil der Darstellung) bildet der Verlauf der Grundfrequenzkontur. Zu erkennen ist, dass im Falle des Neuinformationsfokus (1)(a) die lexikalisch betonte Silbe, die auch den Satzakkzent trägt, keinerlei größere Abweichung von der globalen Grundfrequenzkontur über den gesamten Satz hinweg aufzeigt. Wird dieser Satz in einem Kontext, der einen Kontrast erfordert (1)(b), realisiert (siehe Abbildung 2), so ergibt sich folgendes akustisches Resultat:

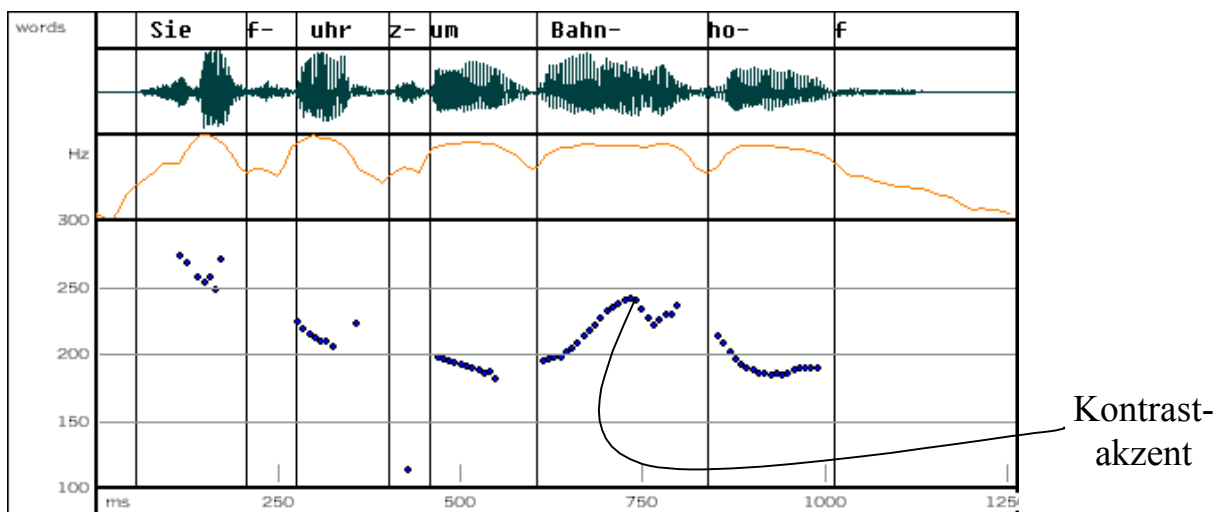


Abb. 2: Ein akustischer Einzelfall des Satzes ‚*Sie fuhr zum Bahnhof.*‘, eingebettet in einen Kontrast auslösenden Kontext (1)(b).

Der Verlauf der Grundfrequenzkontur in Abbildung 2 zeigt im Vergleich zu Abbildung 1 exakt an der Stelle, die auch gleichzeitig den Satzakkzent trägt (an der lexikalisch betonten Silbe von 'Bahnhof' - 'BAHN'), eine deutliche Veränderung. Während in Beispiel (1)(a) keine markante Grundfrequenzbewegung über der Satzakkzenttragenden Silbe auftritt, und die Grundfrequenzkontur über dieser Position in den globalen Grundfrequenzverlauf des gesamten Satzes eingebettet ist, ist ein deutlicher Grundfrequenzanstieg über der kritischen Konstituente in Beispiel (1)(b) zu vermerken. Eine kontrastierte Konstituente wird offensichtlich prosodisch lokal über eine typische Grundfrequenzbewegung markiert.

3.1 Detaillierte Auswertung der prosodischen Daten

Diese ersten Befunde scheinen die Annahme zu gestatten, dass, auch wenn die Akzentpositionen von Kontrast- und Neuinformationsfokus zusammenfallen, ein prosodischer Unterschied erkennbar ist. Der Frage nachgehend, ob dieser beobachtete prosodische Unterschied prototypisch auftritt, haben wir die Sprachaufzeichnungen von 8 deutschen Muttersprachlerinnen in den 2x8 Kontexten analysiert. Die akustischen Analysen bezogen sich auf:

- (1) Unterschiede in der prosodischen Realisierung von Kontrastakkzent versus Default-Satzakkzent.
- (2) Eine mögliche Spezifikation von Kontrast hinsichtlich der prosodischen Domänenbildung: Hat Kontrastierung im Vergleich zur Neuinformations-fokussierung akustisch induzierte Unterschiede hinsichtlich der prosodischen Domänenbildung? Bilden kontrastierte Konstituenten eigene prosodische Domänen?

3.1.1 Neuinformationsakkzent versus Kontrastakkzent

Die prosodischen Analysen zeigen zunächst, dass es einen deutlichen Unterschied in der lokalen vertikalen Auslenkung für den Pitch-Akkzent gibt. Der Kontrastakkzent erhält eine höhere vertikale Auslenkung. Dieser Befund steht im Einklang mit Annahmen über Prominenzrealisierungen (Pheby, 1975), Fokusakkzent-Realisierungen (Schafer et al., 2000; Selkirk, 1995) und Realisierung enger Foki (z.B. Müller, 1998; 1999).

Unsere Daten ergeben einen lokalen Unterschied hinsichtlich der vertikalen Auslenkung der Grundfrequenzkontur in Abhängigkeit von der Einbettung des Zielsatzes in einen Kontext (NIF vs. KF). Die Abbildungen 3 und 4 zeigen den prototypischen Grundfrequenzverlauf (F0), ermittelt von 8 Sprecherinnen für die Beispiele (1)(a) vs. (1)(b):

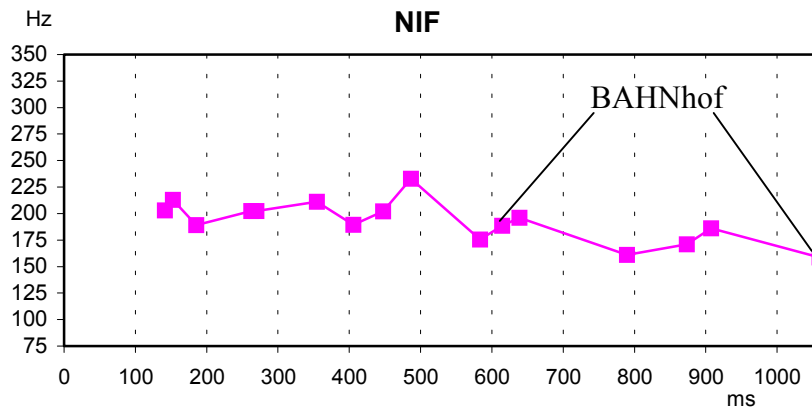


Abb. 3: Prototypische F0-Kontur (8 Sprecherinnen) des Satzes ‚Sie fuhr zum Bahnhof.‘ in Kontext mit Neuinformationsfokus.

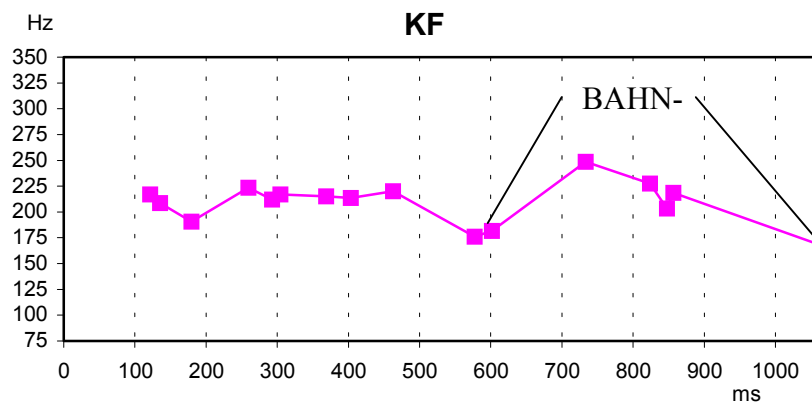


Abb. 4: Prototypische F0-Kontur (8 Sprecherinnen) des Satzes ‚Sie fuhr zum Bahnhof.‘, eingebettet in einen Kontrast auslösenden Kontext.

Ein Kontrastakzent ist hervorgehobener als ein normaler Satzakzent. Die Markierung von Kontrast/Korrektur erfolgt dabei über einen lokalen F0-Anstieg. Bei Neuinformationsfokus ist der Satzakzent dagegen in eine globale fallende Kontur eingebettet. Der Satzakzent ist bei NIF in die globale F0-Deklination, in eine stetig fallende F0-Kontur eingepasst und bleibt damit für den F0-Parameter unmarkiert. Der Kontrastakzent dagegen durchbricht die F0-Deklination und scheint dadurch markierter zu sein als ein durch NIF induzierter Satzakzent.

In den folgenden beiden Graphiken (Abbildung 5 und 6) sind über alle Kontexte (Anekdoten) und über alle Sprecherinnen die lokalen Grundfrequenzwerte vom Offset des unmittelbar vor der Zielkonstituente befindlichen Wortes sowie von Onset, Peak und Offset der betonten Silbe der Zielkonstituente gefolgt vom Onset der nachfolgenden Konstituente/Onset der nachtonigen Silbe mit den dazugehörigen Zeitwerten gemittelt. Die Zielkonstituente ist die Satzakzent tragende bzw. Kontrastakzent tragende Konstituente der Zielsätze.

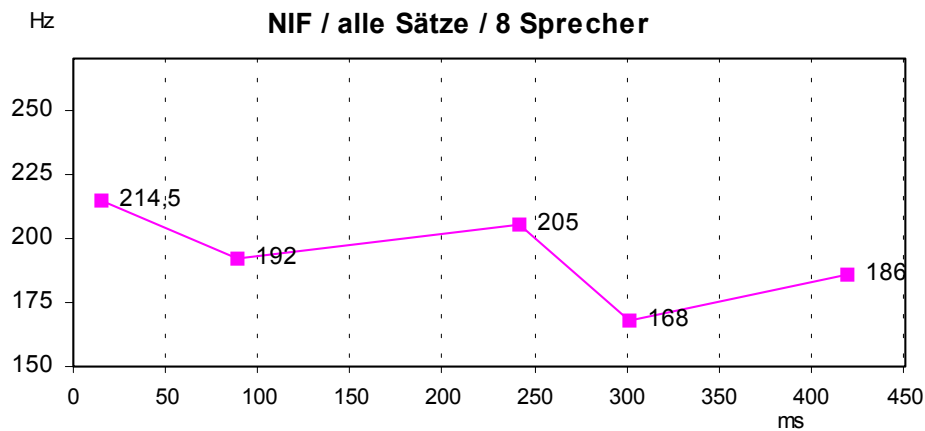


Abb. 5: Prototypische F0-Kontur in Kontext mit Neuinformationsfokus (8 Sprecherinnen, 8 Kontexte).

Der in Abbildung 3 und 4 gezeigte Effekt für einen über 8 Sprecherinnen gemittelten Beispielsatz scheint für das gesamte Satzmaterial (8 Kontexte) stabil zu sein, obwohl die Zielkonstituente variabel hinsichtlich ihrer Stellung im Satz war (satzfinale Stellung bei SVO-Sätzen, nicht-satzfinale Stellung bei SOV-Sätzen).

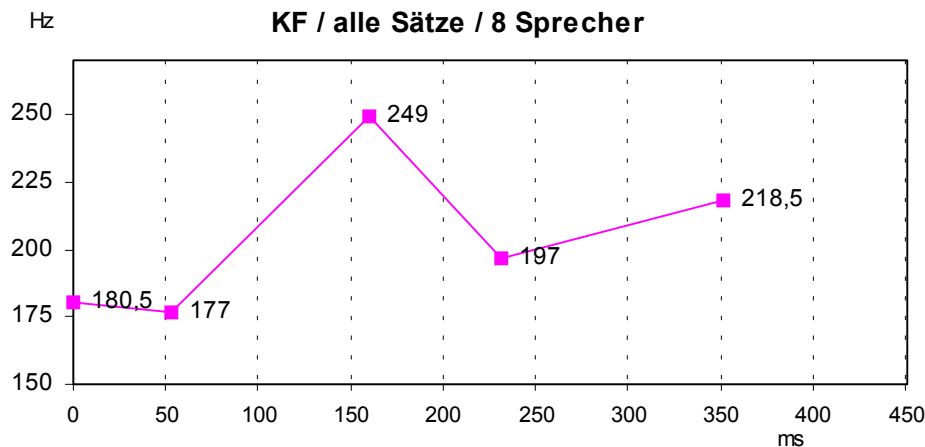


Abb. 6: Prototypische F0-Kontur in Kontext mit Kontrastfokus (8 Sprecherinnen, 8 Kontexte).

Im Falle von Neuinformationsfokus können wir den Satzakkent (lokal) als fallende !H*+L-Sequenz analysieren. Der Offset-F0-Wert der Vorgängerkonstituente weicht nur geringfügig vom F0-Wert der F0-Gipfelposition (Peak) der Zielkonstituente ab. Es ist daher anzunehmen, dass der F0-Peak herabgestuft wurde (*downstep*) und sich in die globale Satzkontur (Deklination) einbettet.

Die linke Flanke des Kontrastakzentes steigt steiler an, und dadurch tritt der F0-Peak früher als beim NIF-Akzent auf. Der Kontrastfokus kann als lokal steigende L+H*-Tonsequenz analysiert werden.

3.1.2 Kontrast/Korrektur und prosodische Domänenbildung

Hier soll der zweiten Fragestellung (s. 3.1 Punkt 2) nach der möglicherweise durch Kontrast induzierten prosodischen Domänenbildung nachgegangen werden.

Die von uns ausgewählten kritischen Messpunkte, um akustisch induzierte Domänengrenzen determinieren zu können, waren die folgenden:

1. Dauerparameter:

- eine präfinale Dehnung derjenigen Konstituente, die unmittelbar links vor derjenigen Konstituente steht, die mit Kontrast- oder Satzakzent assoziiert ist
- eine durch Akzentuierung bedingte Längung der akzenttragenden Silbe
- eine präfinale Dehnung, die mit der rechten Begrenzung derjenigen Konstituente assoziiert ist, die mit Kontrast- oder mit Satzakzent belegt ist
- Pauseninsertion, d.h. phonetische Stille, unmittelbar linksseitig oder rechtsseitig der Zielkonstituente

2. Grundfrequenzparameter:

- ein lokaler Grundfrequenzunterschied im Onset derjenigen Konstituente, die mit dem Satzakzent = Kontrastakzent assoziiert ist
- ein spezifischer Grundfrequenzwert am Offset derselben Konstituente

Auswertung der Dauerparameter:

Zunächst wurde ermittelt, ob sich Dauerunterschiede an der linken Begrenzung der satzakzent-/kontrastakzenttragenden Zielkonstituenten ergaben. Dazu wurde der Ausgangspunkt der Berechnung bei KF (s. Abbildung 6) im Offset der Vorgängerkonstituente im akustischen Signal auf 0 ms festgelegt, da das segmentale Material an dieser kritischen Position phonetisch sehr variabel ist. Die Daten zeigen (s. Tabelle 1), dass für KF im Vergleich zu NIF eine zeitliche Stauchung bis zum Erreichen des F0-Gipfels innerhalb der akzentuierten Silbe auftritt. KF zeigt im Vergleich zu NIF *keine* deutliche lokale Dehnung der akzentuierten Silbe. (vgl. Abbildung 6).

Eine Pauseninsertion zwischen der rechten Begrenzung der Vorgängerkonstituente und dem Onset der Zielkonstituente zeigte zwischen KF und NIF ebenfalls keinen Unterschied (KF: 52 ms; NIF: 73 ms).¹

¹ Die Analyse der präfinalen Dehnung ist in Vorbereitung.

Auswertung der Grundfrequenzparameter:

Die folgenden F₀-Werte (Abbildung 7) zeigen für die linke Begrenzung der Zielkonstituente wiederum einen Unterschied zwischen NIF und KF: Für NIF ist zwischen dem F₀-Offset der Vorgängerkonstituente und dem F₀-Onset der Zielkonstituente ein F₀-Abfall (von 214,5 Hz auf 192 Hz) zu verzeichnen.

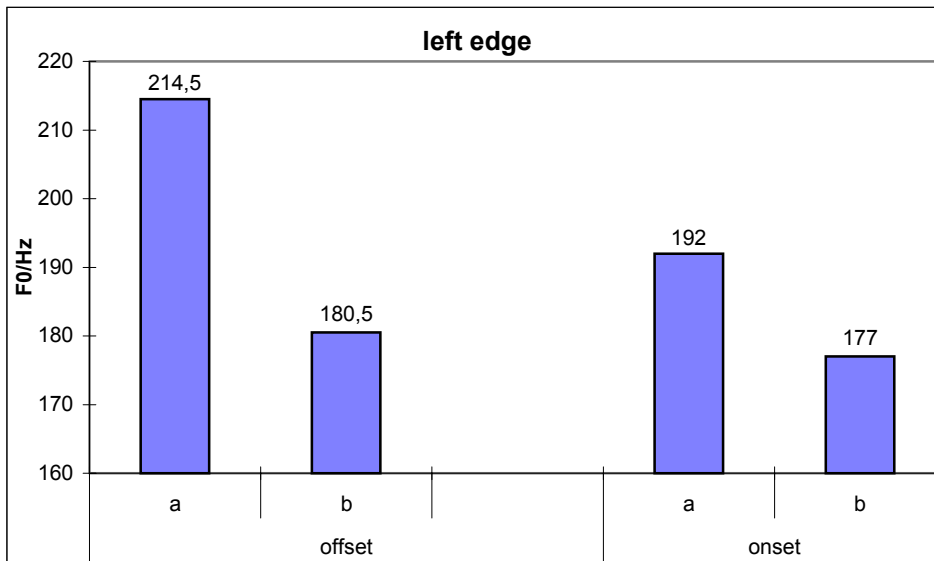


Abb. 7: absolute F₀-Werte in der Umgebung der linksseitigen Begrenzung für NIF (a) und KF (b) der Zielkonstituente für alle Sprecher und alle Kontexte. Offset = rechte Begrenzung der Vorgängerkonstituente; Onset = Beginn der Zielkonstituente

Gemeinsam genommen mit dem unter Abbildung 5 dargestellten F₀-Verlauf für NIF bedeutet dass ein leichtes Absinken der Grundfrequenzkontur linksseitig der Zielkonstituente, dem ein gerinfügiger, durch 'downstep' reduzierter F₀-Gipfel folgt.

Für KF dagegen ist ein starkes Absinken der F₀ am Offset der Vorgängerkonstituente zu beobachten, das sich zum Onset der Zielkonstituente hin abschwächt.

An der kritischen Stelle tritt damit ein Unterschied in der lokalen F₀-Kontur auf: im Fall von NIF wird der Satzakkzent in die globale, stetig fallende F₀-Kontur eingebettet, im Fall von KF wird das Durchbrechen und der veränderte, steigende F₀-Verlauf an dieser Position im Signal bereits vorbereitet.

Die Daten für KF sind jedoch noch nicht als linksseitige bzw. konstituenteninitiale Grenzmarkierung zu bewerten. Hierbei wird lediglich der relativ hohe F₀-Gipfel und die damit korrelierende linke Flanke der L+H*-Tonsequenz vorbereitet. Ähnliches lässt sich bezüglich der F₀ an der rechtsseitigen Begrenzung der Zielkonstituente feststellen.

In Abbildung 8 und 9 werden die rechtsseitigen F₀-Grenzwerte für die Zielkonstituente für NIF (a) und KF (b) gezeigt.

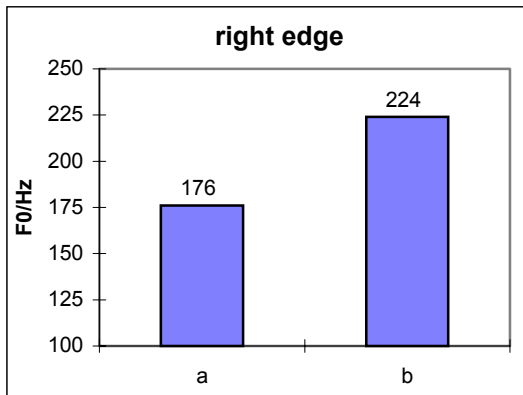


Abb. 8: absolute F0-Werte in der Umgebung der rechtsseitigen Begrenzung für NIF (a) und KF (b) der Zielkonstituente für alle Sprecher und alle Kontexte

In Abbildung 8 sind diese Werte für alle Zielkonstituenten dargestellt. Diese Werte sind hinsichtlich einer möglichen Grenzmarkierung allerdings nur bedingt aussagefähig, da in ihnen auch die Werte für satzfinale Zielkonstituenten enthalten sind. Da alle unsere Zielsätze Deklarative sind, enden diese auch mit einem satzfinalen, tiefen Grenzton (L%). Dieser Grenzton ist Bestandteil der globalen F0-Kontur über den gesamten Satz.

Zunächst scheint es so, als ob für KF im Vergleich zu NIF aufgrund der F0-Differenz ein hoher Grenzton für die rechte Begrenzung einer Intonationsphrase postuliert werden könnte. Die damit normalerweise einhergehenden Dauerparameter wie 'Pauseninsertion' widersprechen jedoch dieser Annahme. Eine Pauseninsertion an der rechten Begrenzung der Zielkonstituente wurde bislang nicht gemessen.

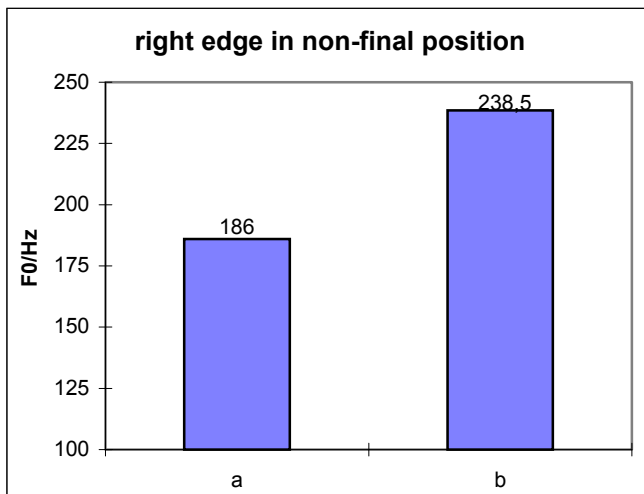


Abb. 9: absolute F0-Werte in der Umgebung der rechtsseitigen Begrenzung für NIF (a) und KF (b) der Zielkonstituenten, die in nicht-satzfinaler Position stehen

Wir analysierten auch diejenigen rechtsseitigen F0-Grenzwerte für die Zielkonstituenten, die in nicht-satzfinaler Position stehen (s. Abbildung 9). Die Unterschiede, die sich hier zwischen

NIF und KF ergeben, scheinen nur bedingt eine Grenzmarkierung für KF (erhöhte F0-Werte im Offset der betonten Silbe der Zielkonstituente) zu markieren. Da die Dauerwerte an dieser Position keine sogenannte 'starke' (Intonationsphrasen-) Grenze signalisieren, gibt es zwei Interpretationsmöglichkeiten dieser Daten:

1. An der rechtsseitigen Position nicht-satzfinaler Zielkonstituenten wird keine prosodische Domänengrenze durch einen hohen Grenzton (H%) markiert. Der höhere rechtsseitige F0-Wert für KF ist bedingt durch den höheren F0-Gipfel-Wert (F0-Peak). Er bleibt zwar auf einem relativ hohen Niveau, ordnet sich jedoch den Bedingungen eines stetig fallenden, deklarativ typischen globalen F0-Verlaufs unter. Dadurch wird dieser F0-Wert in seiner vertikalen Ausdehnung reduziert und kann nicht als Grenzmarkierer identifiziert werden.

2. Eine Intonationsphrasengrenze an dieser Position im Signal scheidet auch aus Gründen des fehlenden folgenden 'F0-Resettings' aus. Der F0-Verlauf der nachfolgenden Konstituenten ordnet sich wiederum in den globalen F0-Verlauf über den gesamten Satz ein. Der höhere rechtsseitige F0-Wert für KF wird zudem nicht soweit abgesenkt, dass er als tieftoniger (L%) Marker einer Intonationsphrasengrenze analysiert werden kann. Möglich wäre noch die Variante eines abgesenkten Phrasentones (L-) für die Markierung einer hierarchisch tiefer angesiedelten prosodischen Domäne wie der 'intermediären oder phonologischen Phrase' (s. Schafer et al., 2000).²

3.2 Schlussfolgerung

Es zeigten sich keine lokalen Effekte, die auf eine initiale und/oder finale durch Kontrast ausgelöste Domänenmarkierung hinweisen. Prosodische Analysen, die die bekannten Indizien für eine Grenzmarkierung betreffen, führten in dem hier vorgestellten Vergleich von Korrekturfokus mit Neuinformationsfokus zu keinem prosodischen Unterschied.

Die akustische Untersuchung und die prosodische Analyse zeigte, dass keine der o.g. Markierungen einen Hinweis darauf gibt, dass unter der Bedingung des Kontrasts eine spezifische prosodische Domäne markiert wird.

Das hier vorgestellte Modell betrifft auf Satzebene zunächst eine Vorhersage des metrischen Musters, das man in Form von metrischen Gittern darstellen kann, wie es für Neuinformationsfokus u.a. von Jacobs (1993) vorgeschlagen wurde. Das Modell bezieht sich dabei

² Wir sind uns bewusst, dass die hier präsentierte Art der **lokalen** Analyse eine Reihe von Fragen offenlässt. Eine umfassende prosodische Analyse unter Einbeziehung der prosodischen Parameter über den **gesamten** Satz hinweg ist derzeit in Vorbereitung (Alter, Mleinek, Umbach, Steube; in Vorbereitung).

auf die Verb-Argument-Struktur. Damit lassen sich die Satzakkente in vollfokussierten und kategorischen Sätzen berechnen. Akzentzuweisung und Prominenzzuweisung in Sprachen mit variabler Wortstellung ist abhängig von der informationsstrukturell festgelegten Fokusdomäne. NIF-Satzakkente treten in der Fokusdomäne auf, die sich wie bei kategorischen Sätzen über den gesamten Satz erstrecken kann.

Die Kontrastierung erfolgt immer mit Bezug auf einen Kontextsatz. Es kann sich bei nur implizierter Kontrastierung auch um eine Proposition handeln, die einen Informationsausschnitt des vom Sprecher als inkorrekt eingestuften *common grounds* vertritt. Sie kann sich auf die ganze Domäne des Satzes erstrecken, auf die Entsprechung eines mittleren NIF, minimalen NIF, aber auch auf deutlich kleinere morphosyntaktische Einheiten.

Bezüglich des (default-) Satzakkentes ist anzunehmen, dass dieser metrischen Gegebenheiten gehorcht. Man kann ihn von der Syntax her und der zugrundeliegenden Verb-Argument-Struktur (Jacobs, 1993) algorithmisch ableiten.

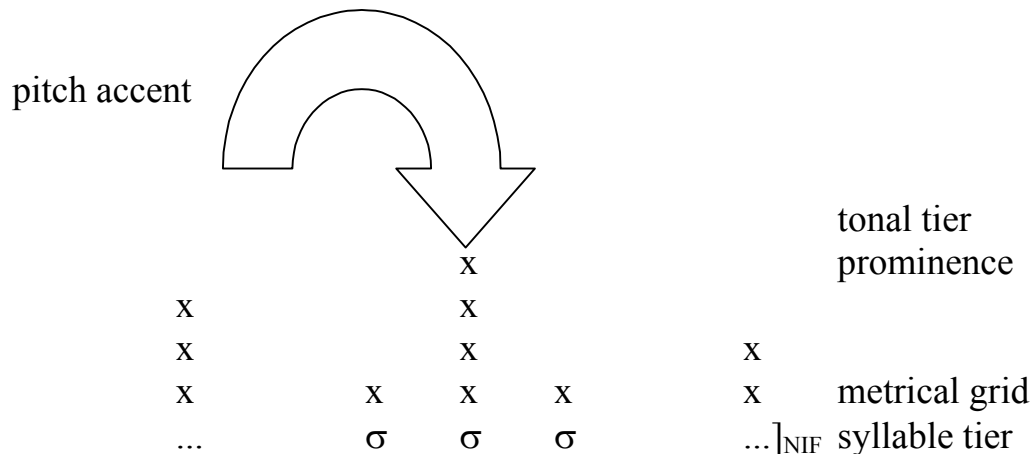


Abb. 10: (Default-) Satzakkent-Zuweisung

Dieser Algorithmus beruht darauf, dass ausgehend von der syntaktischen Struktur des Satzes der Satzakkent derjenigen lexikalisch betonten Silbe des tiefsteingebetteten Verbarguments oder -modifikators zugewiesen wird, die auch den Wort- und/oder Phrasenakzent dafür trägt. In Abbildung 10 wird ein Vorschlag im Rahmen der Autosegmentalen Phonologie (Goldsmith, 1990) vorgestellt, der diesen Gegebenheiten Rechnung trägt. In einer Fokusdomäne ist eine Folge von Silben enthalten. Einige dieser Silben sind lexikalisch betont, d.h., sie sind Träger der Wortbetonung. Bei der Berechnung des metrischen Gitters ist diese Prominenzinformation zu berücksichtigen. Für die Satzakkentzuweisung bei NIF kommen nur diese prominenten Silben in Frage. Unter diesen prominenten Silben wird mittels des von Jacobs vorgeschlagenen Algorithmus diejenige berechnet, die innerhalb der Fokusdomäne bzw. des

Satzes die prominenteste ist. Diese Silbe ist die satzakzenttragende Silbe und wird mit einem pitch-Akzent, d.h., mit einer Tonsequenz assoziiert. Die mit dem Satzакzent assoziierte Tonsequenz hat demzufolge nur direkten Zugriff auf die Fokusdomäne und die in ihr enthaltene Prominenzinformation.

Auf welche (phonologische) Ebene greift nun Kontrastierung im Deutschen zu? Welches ist das probate Mittel, um Kontrast prosodisch zu kodieren?

3.3 Assoziierung mit Kontrast

Kontrastakzente werden kontextabhängig zugewiesen, aber direkt assoziiert mit dem kontrastierten Wort oder mit der kontrastierten Phrase. Sie stimmen dann mit der betonten Silbe des lexikalischen Wortes bzw. der Phrase desjenigen lexikalischen Wortes überein, das auch den Phrasenakzent trägt. Dies sind die kanonischen Fälle, die den Akzentplatzierungsregeln des Lexikons und der Syntax ebenso folgen wie NIF-Akzente (vgl. [*auf dem BAHNhof*]_{KF/NIF}). Der Unterschied zum NIF-Akzent ist aber, dass die lexikalischen und syntaktischen Bedingungen für die Hauptakzentplatzierung nur in dieser KF-Domäne gelten, während der NIF-Akzent sowohl in vollfokussierten als auch in kategorischen Sätzen aus den Wort- und Phrasenakzenten des gesamten Satzes abgeleitet wird.

Daneben kann der Kontrastakzent, der Satzакzent ist, jedoch sowohl die lexikalischen Akzentplatzierungsregeln als auch diejenigen für syntaktische Phrasen überschreiben (Diese Silbe ist mit Großbuchstaben ausgezeichnet.): vgl. [*BahnHOF*]_{KF}, [*AUF*]_{KF} dem Bahnhof. Die Korrektur kann sich sogar auf sub-syllabische Elemente wie den Onset: z.B. in *schlagen* vs. *sagen* oder *saufen* vs. *raufen* beziehen. Selbst der Nucleus wie in *siegen* vs. *sagen* kann kontrastiert werden. Immer ist jedoch die Silbe die phonologische Realisierungsdomäne für den KF (s. Abbildung 11).

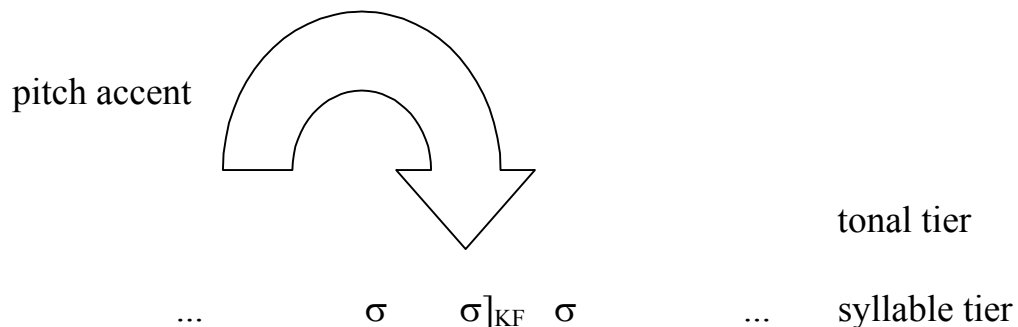


Abb. 11: Ein möglicher Fall von Akzent-Zuweisung bei Kontrastierung

Die mit Kontrast assoziierte Tonsequenz kann somit direkt auf die Silbenebene zugreifen, selbst wenn die betroffene Silbe keinen lexikalischen Akzent (Wortbetonung) trägt. Metrische Information, wie sie u.a. zur Ableitung von Prominenz bei NIF benötigt wird, kann hierbei ausgeblendet werden.

Die Realisierungsdomäne und die KF-Domäne sind jedoch nicht immer identisch, denn die Bedeutungsseite des KF kann sich bei Betroffenheit gebundener Morpheme oder sub-syllabischer Elemente nur relativ zur Wortform realisieren: "SCHLAgen" ist keine korrekte Realisierung des Zeichens "sagen", "BAHNhofen" ist keine korrekte Realisierung des Zeichens "BAHNhof" im Plural. Die Kontrastfokusdomäne ist ein phonologisch-semantisches Korrelat, die Realisierungsdomäne eine prosodische Einheit, die hier mit der Domäne der Silbe spezifiziert werden konnte.

3.4 Zusammenfassung

Anhand der vorliegenden Daten können wir belegen, dass der prosodische Unterschied zwischen NIF und KF von den hier analysierten Sprecherinnen präferiert über einen Unterschied im **Akzenttyp** realisiert wird. Ein KF-induziertes prosodisches Muster weist einen stärker ausgeprägten F₀-Gipfel mit einer vorausgehenden, länger anhaltenden, ansteigenden F₀-Kontur aus (L+H*) auf. Die Zuweisungsdomäne für diesen Akzent ist die Silbe, wobei die zugrundeliegende metrische Struktur für die aus der Wortbetonung abzuleitende Prominenz ausgeblendet werden *kann* (s. 3.3).

Ein durch NIF induzierter Akzent gehorcht den Regularitäten der *default*-Satzakzent-Zuweisung und **mus** metrische Verhältnisse wie lexikalische Wortbetonung berücksichtigen.

Eindeutige Indizien für eine von KF induzierte prosodische Domänenbildung konnten anhand der bislang vorliegenden prosodischen Analysen für das Deutsche nicht gefunden werden.

Die prosodischen Daten verweisen jedoch auf eine eindeutige Präferenz der Kodierung von KF über den Akzenttyp. Ob auch die Akzent*position* einen so deutlichen Unterschied zwischen der prosodischen Realisierung von KF und NIF aufweist, soll Gegenstand weiterführender Untersuchungen sein.

4 Perzeption von Kontrast

Mit der folgenden Studie hatten wir zum Ziel, die prosodischen Unterschiede zwischen der Realisierung von Neuinformationsfokus und Kontrast **perzeptiv** zu testen. Wir gingen dabei der Frage nach, ob entweder (a) für den Hörer vorhergehende Kontextinformation notwendig

ist, um Kontrast zu erkennen oder aber (b) die dem Kontrast zugrunde liegende prosodische Information die Verarbeitung von Kontrast treibt, oder (c) Kontextinformation und prosodische Information gemeinsam einhergehen müssen, um dem Hörer Kontrast-/Neuinformation zu signalisieren.

4.1 Sprachliches Material

16 verschiedene Kontexte sind geschaffen worden, von denen in der Produktionsstudie jeweils 8 Neuinformationsfokus und 8 Kontrast auslösten. Aus diesem Material wurde eine Sprecherin ausgewählt und das von ihr produzierte und von uns aufgezeichnete Material wiederverwendet. Die in Abschnitt 2 beschriebenen Targetsätze wurden kreuzmanipuliert: ein Targetsatz, der von der Sprecherin zunächst in einem neutralen Kontext geäußert worden ist, wurde gegen den von ihr ursprünglich in einem Kontrastkontext geäußerten Satz ausgetauscht und vice versa. Damit erhielten wir schließlich folgendes originales und manipuliertes Sprachmaterial:

- (1) (a) Targetsatz mit Default-Satzakzent in neutralem Kontext (Originalversion)
(b) Targetsatz mit Kontrastakzent in Kontext mit neutralem Kontext (geschnittene Version)
- (2) (a) Targetsatz mit Kontrastakzent in Kontrastkontext (Originalversion)
(b) Targetsatz mit Default-Satzakzent in Kontrastkontext (geschnittene Version)

Jeder Targetsatz wurde zweimal präsentiert, so dass die Versuchspersonen insgesamt 32 Hörproben zu beurteilen hatten. Zwei Hörproben wurden jeweils gegenübergestellt. Sie stammten jeweils entweder aus Gruppe 1 oder Gruppe 2. Die Hörprobe aus Gruppe 1 wurde von derjenigen aus Gruppe 2 getrennt durch Hörproben anderer Kontexte dargeboten.

4.2 Experimentelles Design

Das oben beschriebene Sprachmaterial haben wir in einem Experiment pseudorandomisiert dargeboten. Die insgesamt 16 Kontexte wurden zunächst den Versuchspersonen zusammen mit den Instruktionen und den Fragebögen in schriftlicher Form vorgelegt und nur die sogenannten Hörproben akustisch präsentiert.

Die Fragebögen enthielten zwei Aufgabenstellungen: Die Versuchspersonen hatten erstens Entscheidungen darüber zu treffen, ob sie akustische/prosodische Unterschiede zwischen den zwei jeweils zusammen angebotenen auditorischen Varianten wahrnehmen und zweitens darüber, welche der beiden Varianten besser mit der Interpretation, die durch die ausschließlich

schriftlich dargebotenen Kontexte vorgegeben wurde, zusammenpasste. Insgesamt wurden die Reaktionen von 15 Versuchspersonen mittels Fragebogenanalyse aufgezeichnet. Die im Fragebogen indirekt abgeforderte Fragestellung war demnach, ob die unterschiedliche Semantik von Neuinformationsfokus und Kontrastfokus auch perzeptionsseitig mit bestimmten Akzenttypen assoziiert ist.

4.3 Auswertung der Perzeptionsstudie

Die Frage, ob zwischen der Original- und der *crossgesplichen* Version ein Unterschied hörbar ist, ergab für beide Bedingungen (Kontrast in neutraler Umgebung und NIF in Kontrast-Umgebung) ca. 83%. Dieses Ergebnis wurde wie folgt erzielt: Jeweils 8 Kontexte x 15 Probanden führen zu 120 möglichen korrekten Voten (korrekt heißt hier: prosodische Unterschiede wurden festgestellt). Eine Aufschlüsselung nach Voten ergibt:

- (a) Kontrastakzent in neutraler Umgebung: 99/120 x 'korrekt' = 82,5%
- (b) neutraler Satzakzent in Kontrast-Umgebung: 100/120 x 'korrekt' = 83,3%

Die Auswertung der Frage, welche der beiden Varianten besser mit der Kontextinformation zusammenpasste, ergab ein ziemlich heterogenes Bild. Zunächst sind wiederum 120 korrekte Voten möglich, wobei korrekt in diesem Falle bedeutet, dass die Probanden die Original- der kreuzmanipulierten Version vorzogen. Von diesen 120 Voten wurden die in Frage 1 ermittelten Voten abgezogen, bei denen kein Unterschied wahrgenommen wurde. Zusätzlich wurde ein Kontext ganz aus der Wertung herausgenommen, da er den Trend konterkarierte.³ So erhielten wir schließlich als Ausgangsbasis 86 mögliche korrekte Voten bei Kontrast in neutraler Umgebung sowie 89 mögliche korrekte Voten bei NIF-Akzent in Kontrast-Umgebung (korrekt heißt hier: der Mismatch wurde erkannt). Das Ergebnis nach Voten und in Prozent liegt wie folgt vor:

- (a) bei Kontrastakzent in neutraler Umgebung: **47/86 x 'korrekt' = 55%**
 - (a1) Mismatch wurde nicht erkannt: 23/86 x 'inkorrekt' = 27%
 - (a2) Enthaltung: 16/86 x 'Enthaltung' = 19%
- (b) bei NIF-Akzent in Kontrast-Umgebung: **73/89 x 'korrekt' = 82%**
 - (b1) Mismatch wurde nicht erkannt: 3/89 x 'inkorrekt' = 3,4%
 - (b2) Enthaltung: 13/89 x 'Enthaltung' = 15%

³ Tatsächlich war der Kontext anders konstruiert. Er basierte auf einem Funktionsverbgefüge und war insofern nicht vergleichbar mit den anderen Kontexten.

Die folgende Graphik verdeutlicht dieses Ergebnis nochmals:

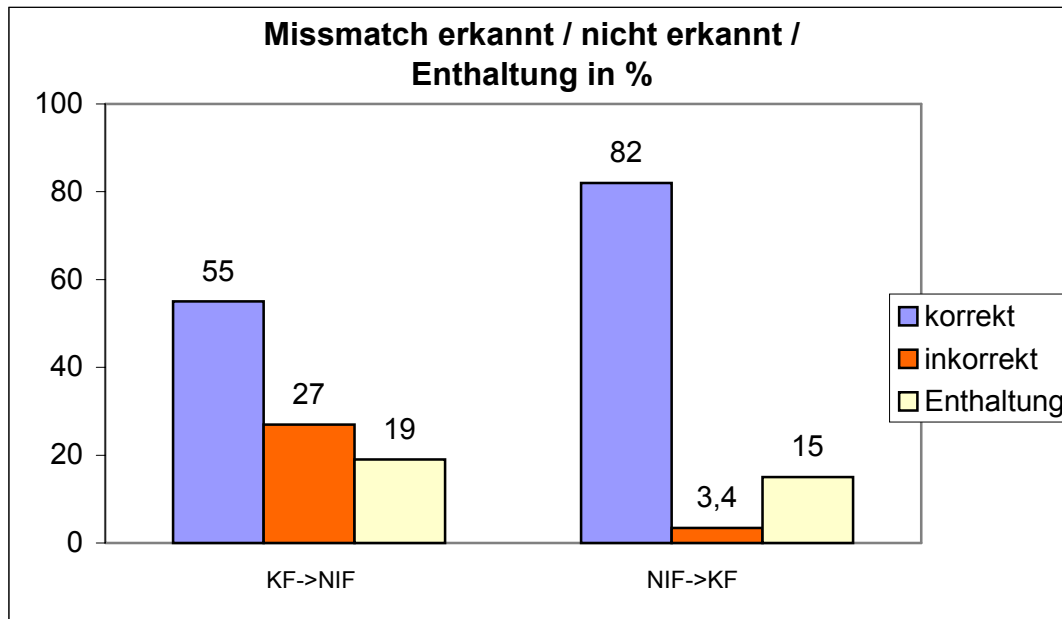


Abb. 12: Positive und negative Mismatcherkennungsraten sowie Enthaltungen in Prozent

Im folgenden werden beide Bedingungen (Kontrastakzent in neutraler Umgebung (= KF → NIF) und Default-Satzakzent in Kontrastumgebung (= NIF → KF)) getrennt betrachtet, da deutliche Unterschiede bei der Bewertung zwischen den beiden kreuzgeschnittenen Versionen des Materials auftraten.

4.3.1 Kontrastakzent in neutraler Umgebung

Die Perzeptionsstudie ergab, dass bei Darbietung eines Kontextes, der für Neuinformationsfokus steht, in den jedoch ein prosodisch markierter Kontrastfokus eingebettet ist, der prosodische Mismatch nur marginal erkannt wird. Die Erkennungsrate des Mismatches liegt bei 55%, und das entspricht der Zufallsrate. Mit anderen Worten, man kann nichts darüber aussagen, ob die Probanden den Mismatch erkannt haben oder nicht. Einigen Aufschluss kann noch die Auswertung der einzelnen Kontexte bringen. Siehe dazu Abbildung 13:

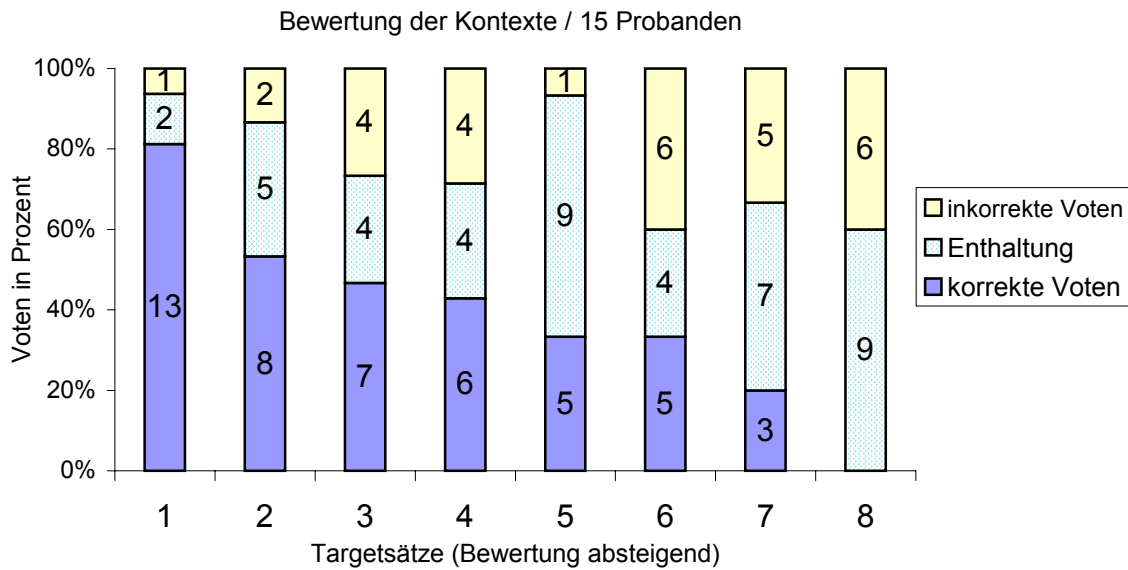


Abb. 13: Auswertung der Kontexte nach absoluten Voten bei KF->NIF

Einige unserer Kontexte scheinen geeignet zu sein, dass Probanden den prosodischen Missmatch deutlich erkennen (z.B. Kontext 1), andere wiederum scheinen offenbar schwieriger zu beurteilen zu sein. Damit ergeben sich ähnliche Probleme, wie sie in Allbritton et al. (1996) berichtet wurden. In den von diesen Autoren durchgeführten Studien zur Sprachproduktion und -perzeption wurden ebenfalls Schwierigkeiten bei der perzeptiven Verarbeitung von Kontextinformation beobachtet.

Die vorläufige Schlussfolgerung zur gesamten Datenlage ist, dass ein Mehr an prosodischer Parameterausprägung die Verarbeitung von Neuinformationsfokus nicht stört. Dies scheint zu bedeuten, dass eine lokale Hyperartikulation, Hervorhebung, oder Überbetonung an der Position des Satzakzentes von den Probanden akzeptiert wird.

Zwei Szenarien sind vorstellbar:

1. Die prosodische Information wird ausgeblendet, der Unterschied im Akzenttyp wird beispielsweise als sprechertypische Variation interpretiert. Der semantischen Interpretation wird dadurch der Vorzug gegeben; die prosodische Information dagegen scheint bei der Verarbeitung unterspezifiziert zu sein. Dies wäre in Einklang mit Hypothese (a) in 4.

2. Der Kontrastakzent wird während der Verarbeitung verwertet und der Kontext entsprechend interpretatorisch angepasst. Eine Besonderheit bei Kontrast/Korrektur ist, dass sie unmittelbar auf geeignete Situationen appliziert werden können. Ein NIF-induzierender Kontext könnte daher eine Uminterpretation in einen KF-Kontext erfahren haben. Dies ist aber inhaltlich kaum zu verhindern gewesen. Wir haben versucht, die neutralen Kontexte so zu konstruieren, dass sie in jedem Fall neutral zu interpretieren sind, dennoch hat sich bei Befragungen

herausgestellt, dass auch diese Kontexte eine mögliche kontrastive Interpretation haben können. Die prosodische Information wird demnach nicht ausgeblendet. Semantische und prosodische Information scheinen in diesem Fall eng miteinander zu interagieren, was Hypothese (c) in 4 entspricht.

4.3.2 Default-Satzakzent in Kontrast evozierendem Kontext

Wie aber sieht es nun aus, wenn kontextuell bedingt ein Kontrast erwartet wird, jedoch im Zielsatz die eher schwach ausgeprägten prosodischen Parameter, die für NIF typisch sind, kodiert sind. Hier liegt die Erkennungsrate des prosodischen Missklangs deutlich höher, unserer Studie entsprechend bei 82%.

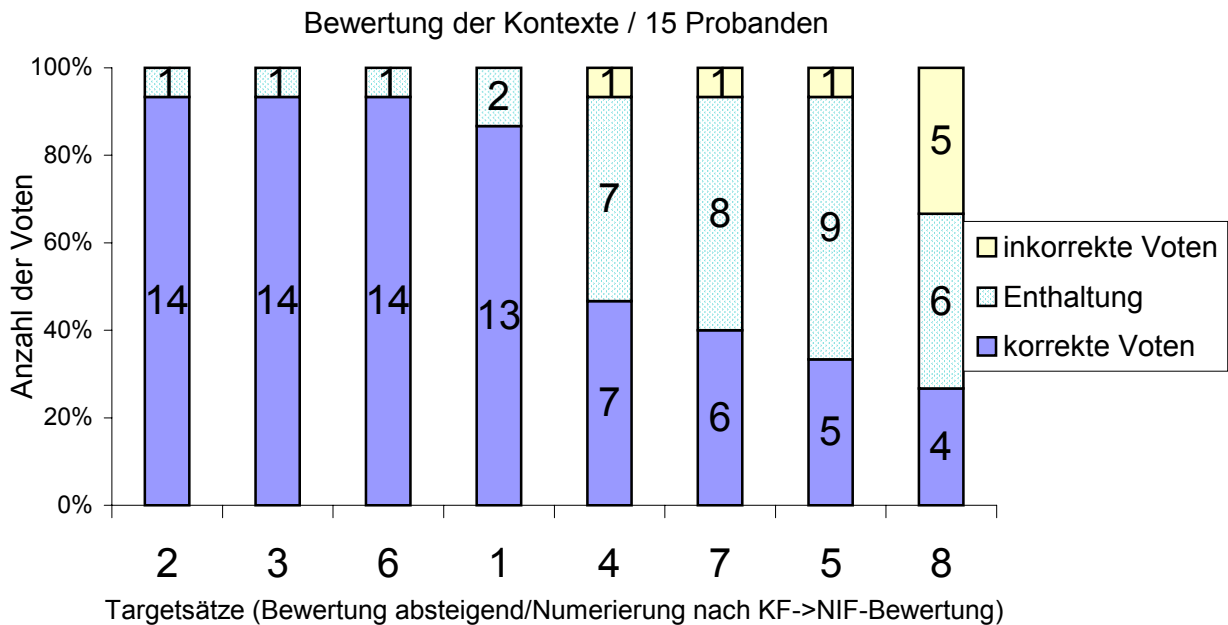


Abb. 14 Auswertung der Kontexte nach absoluten Stimmen bei NIF->KF

Einige Realisierungen (die Kontexte 2, 3, 6, 1) werden sogar von allen Probanden einheitlich abgelehnt (siehe dazu Abbildung 14). Eine fehlende prosodische Markierung für einen erwarteten KF-Akzent wird demnach als unpassend betrachtet. Dies kann zweierlei bedeuten:

1. Der KF-Akzent unterscheidet sich perzeptiv vom NIF-Akzent. Der mit Kontrastfokus einhergehende Akzent wird demnach erwartet, angeboten wird dagegen nur ein neutraler Satzakzent. Dies wird als prosodischer Fehler bemerkt. Prosodische Information ist demnach wichtig für die Signalisierung von Kontrast. Dies entspricht Hypothese (b) in 4.

2. Die semantische Interpretation hat zwar den Vorrang, aber die Verarbeitung eines Kontrastkontextes gilt erst dann als gelungen, wenn die prosodischen Parameter ebenfalls optimal an diesen Kontext angepasst sind, was Hypothese (c) in 4 entspricht.

Damit bleibt die Frage offen, welche der sprachlichen Ebenen - Semantik oder Prosodie - die relevantere für die Verarbeitung von Kontexten ist, oder ob beide Ebenen miteinander interagieren. Man kann der Frage nachgehen, indem man eine direkt auf die kognitiven Verarbeitungsprozesse zugreifende, zeitlich hochauflösende Methode verwendet wie bspw. neurophysiologische Untersuchungen mittels ereigniskorrelierter Hirnpotentiale (EKP). Damit ließe sich u.U. die zeitliche Abfolge von Verarbeitungsschritten (semantische Verarbeitung *vor* prosodischer Verarbeitung oder umgekehrt) klären.

Eine weitere Möglichkeit wäre, die Frage zu konkretisieren und zu ermitteln, ob es Verarbeitungsunterschiede hinsichtlich 'überflüssiger' oder 'fehlender' Akzentmarkierung gibt. Ein erster Schritt in diese Richtung wurde von Hruska et al. (2000; 2001) getan. Die auf neurophysiologischen Untersuchungen beruhenden Daten von Hruska et al. (2001) legen zunächst nahe, dass eine inkorrekte **enge** NIF-Akzentuierung⁴ (wie im Beispiel: Was verspricht Peter Anna zu tun? Peter verspricht **ANna** zu entlasten und das Büro zu putzen.) vom Hörer sofort durch eine parietale Negativierung im EKP angezeigt wird. Dies ist jedoch nicht der Fall für den sogenannten 'überflüssigen' Akzent auf *ANna*. Nur ein nicht vorhandener Fokusakzent (der auf *zu entLASTen* zu erwarten wäre) löst eine parietale Negativierung im EKP aus, während ein zusätzlicher Akzent nicht als prosodische Verletzung interpretiert wird.

5 Schlussfolgerung

Akustisch messbare Unterschiede müssen nicht zwangsläufig auch mit perzeptiven kategorisierbaren Unterschieden einhergehen. Spezifische Grundfrequenzunterschiede im akustischen Signal, wie sie hier für NIF- und KF-Akzentrealisierung determiniert wurden, müssen nicht notwendigerweise perzeptive Relevanz erlangen. Das auditorische und kognitive System ist zu einer deutlich flexibleren Wahrnehmungsbreite in der Lage. Prosodische Anomalien können während der Perzeption ignoriert oder unterdrückt werden, sobald die Aufgabenstellung dies erfordert. Betrachtet man jedoch unsere Daten, ergibt sich, dass Kontrastierung mit einer spezifischen Art von prosodischer Ausbuchstabierung einher gehen kann und dies auch akustisch messbar ist. Ob es aber perzeptiv relevant ist oder nicht, lässt die Studie offen. Zu ähnlichen Befunden kommt die Studie von Studdert-Kennedy & Hadding, 1973.

⁴ In einer Experimentalreihe wurde mit Frage-Antwort-Paaren gearbeitet. Dabei wurden w-Fragen mit prosodisch korrekter und inkorrekt Akzentuierung in den Antwortsätzen operiert.

Literatur

- Allbritton, D.W., McKoon, G. & Ratcliff, R. (1996): Reliability of Prosodic cues for resolving syntactic ambiguities. *Journal of Experimental Psychology: Learning, Memory And Cognition*, 22 (3):714-735.
- Alter, K. (1997): Russian Prosody: Phrasing and tonal structure. In: *Michigan Slavic Publications, Proceedings of Formal Approaches to Slavic Languages V*, Indiana. pp. 1-17.
- Alter, K., Astésano, C., & Besson, M. (2001): Influence des modalités intonatives sur le traitement sémantique: une étude électrophysiologique. In: Cavé, Ch., Guaitella, I. & S. Santi (eds.). *Orality and Gestures. Interactions et comportements multimodaux dans la communication*. l'Harmattan. Paris. 411-415.
- Bolinger, D.L. (1961): Contrastive accent and contrastive stress. *Language* 37: 83-96.
- Cooper, W.E. et al. (1985): Acoustical aspects of contrastive stress in question answer contexts. *JASA* 77:2141-2156.
- Dietrich, R. (1990): Zu Form und Bedeutung der Kontrastintonation im Deutschen, in: *Linguistische Berichte* 129, Westdeutscher Verlag. 415-430.
- Garding, E. (1981): Contrastive prosody: a model and its application". *Studia Linguistica* 35:146-165.
- Goldsmith, J. (1990): *Autosegmental and Metrical Phonology*. Oxford and Cambridge, MA: Basil Blackwell.
- Hruska, C., Steinhauer, K., Alter, K. & Steube, A. (2000): ERP Effects of Sentence Accents and Violation of the Information Structure. Poster presented at the conference 'LabPhon 7', Nijmegen.
- Hruska, C., Alter, K., Steinhauer, K. and Friederici, A.D. (2001): Perception of prosodic cues depends on context: Evidence from ERP studies. 8th Annual Meeting of the Cognitive Neuroscience Society, New York, USA, March 2001. *Journal of Cognitive Neuroscience, Supplement*, 159.
- Hruska, C, Alter, K., Steinhauer, K., & Steube, A. (2001): Misleading dialogues: Human's brain reaction to prosodic information. In: Cavé, Ch., Guaitella, I. & S. Santi (eds.). *Orality and Gestures. Interactions et comportements multimodaux dans la communication*. l'Harmattan. Paris. 425-430.
- Jacobs, (1993): Integration. in: Reis M. (Hrsg.). *Wortstellung und Informationsstruktur*. Linguistische Arbeiten 306. Niemeyer. Tübingen. 63-116.
- Mehlhorn, G./Zybatow, G. (2000): Alte und neue Hüte der russischen Informationsstruktur. In: Dölling, J./Pechmann, T. (Hrsg.): *Prosodie - Struktur - Interpretation*. (= *Linguistische Arbeitsberichte*; 74). Universität Leipzig, S. 25-42.
- Mehlhorn, G. (2001): *Kontrastierte Konstituenten im Russischen: Experimentelle Untersuchungen zur Informationsstruktur*. Phil. Diss., Leipzig
- Müller, K. (1998): *German Focus Particles and their Influence on Intonation*. Master Thesis. University of Stuttgart.
- Müller, K. (1999): *German Focus Particles and Intonation*, Proceedings of the 14th International Conference on Phoenetic Sciences, San Francisco, CA, 1521-1524.
- Pheby, J. (1975): *Intonation und Grammatik im Deutschen*. Sammlung Akademie-Verlag 19 'Sprache' Berlin.
- Schafer, A., Carlson, K., Clifton, C. & Frazier, L. (2000): Focus and the Interpretation of Pitch Accents: Disambiguating Embedded Questions. *Language and Speech*, 43 (1), 75-105.
- Selkirk, E. (1995): Sentence Prosody: Intonation, Stress and phrasing. In Goldsmith, J. A., ed.: *The Handbook of Phonological Theory*. Oxford, UK: Blackwell Publishers Ltd., pp. 550-570.
- Steedman, M.J. & Altman, G.T.M. (1989): Ambiguity in context: A reply. *Language and Cognitive Processes*, 4, 105-122.
- Steube, A. (2000): Ein kognitionswissenschaftlich basiertes Modell für die Informationsstrukturierung (in Anwendung auf das Deutsche). In: J. Bayer, Ch. Römer (Hrsg.): *Von der Philologie zur Grammatiktheorie*. Tübingen: Niemeyer, 213-238
- Steube, A. (im Druck): Correction by Contrastive Focus. In: *Theoretical Linguistics*
- Studdert-Kennedy, M., & Hadding, K. (1973): Auditory and linguistic processes in the perception of intonation contours. *Language and Speech*, 16, 293-313.
- Zybatow, G./Mehlhorn, G. (2000): Experimental Evidence for Focus Structure in Russian. In: King, T. H./Sekerina, I. A.: *Annual Workshop on Formal Approaches to Slavic Linguistics. The Philadelphia Meeting 1999*. Michigan Slavic Publications: Ann Arbor, pp. 414-434.