

DISAMBIGUIERUNG GLEICHER WORTFOLGEN IM DEUTSCHEN
DURCH PROSODISCHE VERÄNDERUNGEN

STEPHAN FORSTNER

Institut für Phonetik und Sprachliche Kommunikation
der Universität München
Schellingstr. 3, 8000 München 40, F. R. G.

ABSTRACT

Das Ausgangsmaterial für die Untersuchung bilden Äußerungspaare von vier männlichen Sprechern des Deutschen, diese Satzpaare weisen zum einen ein- und dieselbe Wortfolge, zum anderen unterschiedliche syntaktische Strukturen auf. Es wird untersucht, ob und in welcher Weise die phonetischen Parameter der Dauer, des FO-Bereichs und der Intensität einzeln und in Abhängigkeit voneinander eine disambiguierende Funktion bei oberflächenstrukturell ambigen Wortfolgen erfüllen.

EINFÜHRUNG

Wenn man davon ausgeht, daß gleiche Wortfolgen in realen Kommunikationssituationen prosodisch verschieden realisiert werden, und daß diese unterschiedlichen Realisationen auch unterschiedliche kommunikative Funktionen erfüllen, so stellt sich die Frage, durch welche prosodischen Veränderungen diese Effekte zustande kommen. In der vorliegenden Untersuchung soll gezeigt werden, inwiefern die phonetischen Parameter FO, Dauer und Intensität bei der Disambiguierung solcher oberflächenstrukturell ambiger Sequenzen eine Rolle spielen.

Zu diesem Zweck wurden die Produktionsdaten mehrerer Sprecher erhoben und interpretiert. Sie sollen die Grundlage für experimentalphonetische Untersuchungen bilden, durch die geklärt werden soll, welche spezielle Funktion jeder Parameter für sich und in Relation zu den anderen Parametern erfüllt.

DAS AUSGANGSMATERIAL

Es wurde nach Material gesucht, das neben oberflächenstruktureller Ambiguität auch die Kriterien der Rekonstruierbarkeit natürlicher Kontexte, der entsprechenden Diskriminierbarkeit dieser Kontexte und der phonetischen Tauglichkeit zur akustischen Analyse und Manipulation natürlich produzierter Äußerungen erfüllt.

Aus einer Reihe von Satzpaaren verschiedenen Typs wurde gemäß der oben genannten Kriterien das folgende Satzpaar ausgewählt:

- a) Ida meinte: "Nina hat angerufen".
- b) "Ida" meinte Nina, "hat angerufen".

Für Äußerungen dieses Typs sind natürliche Kontexte durchaus darstellbar. Die Disambiguierung erfüllt eine klare kommunikative Funktion (etwa zur Beseitigung eines Mißverständnisses). Dementsprechend läßt sich auch die experimentalphonetische Aufgabenstellung formulieren (z. B. als Frage

nach dem Namen des Sprechers, der etwas sagte oder meinte).

PRODUKTION DES AUSGANGSMATERIALS

Bei der Realisation des Ausgangsmaterials erhielten vier männliche Sprecher, deren normales Sprachverhalten keine dialektale Färbung aufweist, die Anweisung, das Satzpaar mindestens drei mal zu äußern. Der den Äußerungen zugrunde liegende Text war jeweils in eine Kontextbeschreibung eingebettet, aus der hervorging, daß die Sprecher eine Situation rekonstruieren sollten, in der ein Mißverständnis aufzuklären war. Ein imaginärer Gesprächspartner sollte davon überzeugt werden, daß Ida gemeint hat. Nina (und nicht Karin) habe angerufen. Analog dazu sind Kontextbeschreibung und Realisation von b) zu sehen. Zwangsläufig wurde dabei von allen Sprechern ein Kontrastakzent gesetzt.

ANALYSE DER DAUERVERHÄLTNISSSE, DES FO-VERLAUFS UND -ONSETS UND DER INTENSITÄT

Das in einem schalltoten Raum auf Analogband aufgenommene Material wurde in eine digitale Kopie überführt, das resultierende digitale Oszillogramm am Bildschirm unter auditiver Kontrolle in zweifacher Weise segmentiert. Zum einen wurden die Segmente jeweils an den einzelnen Lautgrenzen definiert, und zwar bis zum Beginn des letzten Teils der Äußerung, also bis "hat angerufen". Die Dauern dieser Segmente wurden in Millisekunden gemessen. Zum anderen wurden die stimmhaften Passagen derselben Äußerungsteile periodenweise segmentiert, um die Hz- und db-Werte zu ermitteln.

In den folgenden acht Abbildungen sind die den ermittelten Zahlenwerten entsprechenden FO-Verläufe skizziert.

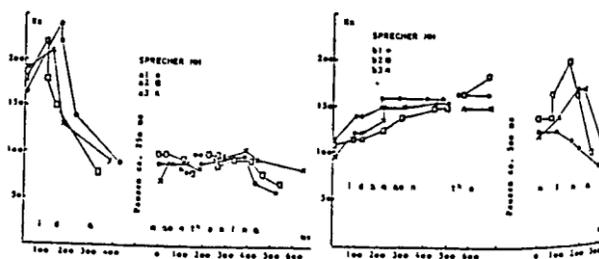


Abb. 1

Abb. 2

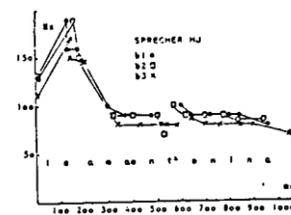


Abb. 3

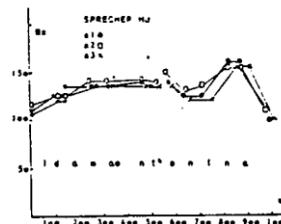


Abb. 4

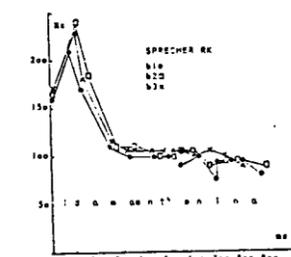


Abb. 5

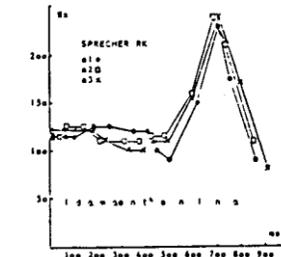


Abb. 6

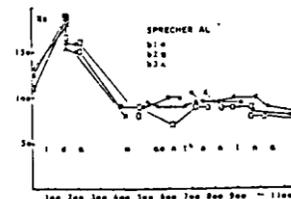


Abb. 7

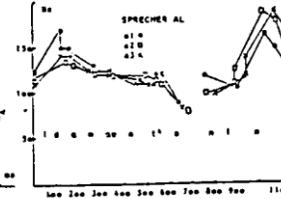


Abb. 8

1. Dauern

Für die Interpretation der Daten war hinsichtlich aller drei phonetischen Parameter in erster Linie die unmittelbare Umgebung der Phrasierungsgrenzen interessant, also der Schwalaut und die anschließende Sequenz /nina/ in den a-Versionen, die Sequenzen /ida/ und /nina/ in den b-Versionen.

Die vier Sprecher wiesen durchaus Unterschiede in der phonetischen Form ihres Disambiguierungsverhaltens auf, dennoch sind einige gemeinsame Regularitäten festzustellen.

Bei den Sprechern AL und MH ist zu sehen, daß an den Phrasierungsgrenzen zumindest partiell Pausen gesetzt werden, während die Sprecher HJ und RK in allen Fällen ohne den Einsatz von Pausen disambiguieren.

Es läßt sich also feststellen, daß die Pausensetzung als Disambiguierungsmittel zweifellos eine entscheidende Funktion übernehmen kann, zugleich stellt sie jedoch keinen zwingend notwendigen Parameter der Disambiguierung dar.

An den Stellen, an denen Pausen gesetzt werden, ist das sogenannte 'pre-pausal lengthening', also eine längere Vokaldauer zu beobachten. In den a-Versionen ist hiervon der Schwalaut betroffen. Beim Sprecher AL liegen die Werte hier um durchschnittlich ca. 60 ms über denen der b-Versionen, beim Sprecher MH um durchschnittlich ca. 100 ms.

In den b-Versionen ist pre-pausal lengthening jeweils im a von "Ida" und "Nina" zu beobachten. Beim Sprecher MH liegen die

Werte des a von "Ida" um durchschnittlich 110 ms über denen der a-Versionen. Nach dem a in "Nina" wird nur in einem Fall eine Pause gesetzt (b3). Der Wert liegt hier um mehr als 100 ms über den Werten von b1 und b2 und auch deutlich über den entsprechenden Werten der a-Versionen. Bei Sprecher AL wird in den b-Versionen nach "Ida" nur in einem Fall eine Pause gesetzt (b3), ohne daß sich der betreffende Wert wesentlich von denen der beiden anderen b-Realisationen unterscheidet. Bei diesem Sprecher zeigt sich jedoch im nachfolgenden Nasal von "meinte" (b1 und b2) ein deutlicher Glottalisierungseffekt, der sowohl auditiv als auch am digitalen Oszillogramm verifiziert werden kann. Die betreffenden Nasalsegmente weisen Dauern von 111 und 155 ms auf gegenüber 52 ms in b3. Diese Beobachtung legt die Vermutung nahe, daß Glottalisierung an Phrasierungsgrenzen anstelle von Pausensetzung und pre-pausal lengthening eine disambiguierende Funktion übernehmen kann. Dies wäre systematisch zu untersuchen. Auch nach dem finalen Vokal von "Nina" ist bei AL in den b-Versionen nur in einem Fall eine Pause vorhanden, ohne daß sich der entsprechende Wert von denen der beiden anderen b-Realisationen unterscheidet. Auch hier ist nachfolgende Glottalisierung zu erwarten.

Bei den Sprechern HJ und RK ist pre-pausal lengthening nicht festzustellen. Die betreffenden Werte der a- und b-Versionen unterscheiden sich nur in geringem Maße voneinander.

Eine weitere Auffälligkeit in den Dauerbeziehungen ist in Zusammenhang mit der Akzentsetzung zu sehen. Die Hauptakzente liegen durchwegs auf der jeweils ersten Silbe von "Nina" in den a-Versionen und von "Ida" in den b-Versionen. Es zeigt sich, daß die akzenttragenden Silben nahezu durchgehend höhere Werte aufweisen als die direkt vergleichbaren Werte der jeweils anderen Version. Die einzige Ausnahme bilden die a-Realisationen von "Nina" des Sprechers MH. Die Dauerwerte für die erste Silbe unterscheiden sich hier nicht wesentlich von den Werten der b-Versionen. Es ist zu vermuten, daß dies mit dem sprecher-spezifischen Rhythmus und Tempo zusammenhängt, die vorhergehenden Pausen weisen immerhin eine relativ hohe Dauer von ca. 500 ms auf.

2. FO-Onset und -verlauf

Bei den Sprechern HJ und AL weist der Schwa-Laut in den a-Versionen durchgehend ein Absinken der Tonhöhe zwischen 20 und 40 Hz auf, während die vergleichbaren FO-Werte in den b-Versionen progressiv verlaufen oder gering abfallen. Zugleich liegt der FO-Onset der a-Versionen um bis zu 50 Hz höher als der der b-Versionen. Allgemein ist zu sagen daß die Schwa-Laute der b-Versionen aus einer nur sehr begrenzten Anzahl von Stimntonperioden bestehen, an einigen Stellen war es überhaupt nicht möglich, periodenweise zu segmentieren.

Bei den Sprechern RK und MH verläuft der Schwa-Laut in den a-Versionen progressiv oder zeigt einen Anstieg, während in den b-Versionen vor allem Progressivität (in einem Fall ein leichter Anstieg um 10 Hz) bei durchwegs niedrigerem Onset zu beobachten ist.

Somit läßt sich bereits ein erstes wichtiges Ergebnis aus der Erhebung der Produktionsdaten formulieren: Sowohl Level- als auch Fall- und Rise-Bewegungen können an gleicher Stelle als Mittel zur Disambiguierung eingesetzt werden. Als wesentlich konsistenterer Parameter erweist sich hier der FO-Onset.

Die akzenttragende Silbe von "Nina" weist im vokalischen Bereich bei allen Sprechern durchgehend einen deutlichen Anstieg um bis zu 90 Hz auf, während die vergleichbaren Silben in den b-Versionen bei ebenfalls deutlich niedrigerem Onset in allen Fällen progredient verlaufen.

Die zweite Silbe von "Nina" zeigt in den a-Versionen ein ebenso deutliches FO-Absinken um bis zu 100 Hz. In den b-Versionen ist dagegen nur ein leichtes Absinken um durchschnittlich ca. 10 Hz oder ein progredienter Verlauf festzustellen, auch hier bei durchgehend niedrigerem Onset.

Auch die akzenttragenden Silben von "Ida" in den b-Versionen zeigen eine deutliche Rise-Bewegung, während in den vergleichbaren a-Versionen Level- sowie leichte Rise- und leichte Rise-Fall-Bewegungen zu sehen sind. Ebenso weist die zweite Silbe von "Ida" in allen b-Versionen ein deutliches Absinken auf. In den a-Versionen findet sich dagegen in erster Linie Progredienz, aber auch ein leichtes Absinken und ein leichter Anstieg des Stimmtones.

3. Intensität

Hier sind in erster Linie die an den akzenttragenden Silben gemessenen db-Werte sowie jene an den damit direkt vergleichbaren Stellen der jeweils anderen Versionen interessant, also die jeweils ersten Silben von "Ida" und "Nina". Dabei war zu erwarten, daß die db-Werte der hauptakzenttragenden Silben wesentlich höher seien als die der nicht akzentuierten. (Ein Unterschied von ca. 6 db ist mit etwa doppelt so hoher bzw. niedriger Lautstärke gleichzusetzen.

Was die erste Silbe von "Nina" betrifft, so hat sich dies durch die erhobenen Daten voll bestätigt, während insbesondere bei den Sprechern MH und HJ in der ersten Silbe von "Ida" das erwartete Ergebnis nicht vorliegt. Dieser Umstand deutet darauf hin, daß bei diesen Sprechern im initialen Teil der Äußerungen wesentlich stärker mit FO- und Dauerverhältnissen disambiguiert wird.

ZUSAMMENFASSUNG

Die erhobenen Produktionsdaten von vier Sprechern zeigen in deutlicher Weise, daß Dauerverhältnisse, FO-Bewegungen und FO-Onset sowie die Intensität phonetische Parameter zur Disambiguierung strukturell ambiger Satzpaare sind. Daher ist auf der Basis der hier ermittelten Daten eine Grundlage für die Generierung von Teststimuli bzw. Testkontinua gegeben, bei denen die genannten Parameterwerte so verändert werden sollen, daß es möglich wird, genauer zu untersuchen, unter welchen Bedingungen Ambiguität bzw. Disambiguität resultiert.