

Segmentierung sprachlicher Information

Stefanie Kruck & Kai Alter

Max-Planck-Institut für neuropsychologische Forschung, Leipzig

Für die Segmentierung eines kontinuierlichen Sprachschalls können prosodische Parameter als Anhaltspunkte dienen. Unter anderem sind es in diesem Zusammenhang (1) Domänengrenzen und (2) Prominenzmerkmale bzw. Betonungen, die genutzt werden, um das akustische Signal zu modifizieren und zu strukturieren. Folgend werden Beispiele präsentiert, die beide Aspekte dieser Strukturierungsmöglichkeiten betreffen. Die darauf basierende Studie befasst sich mit der Verarbeitung von Phonemkategorien, Silbengrenzen und Betonungen mit besonderem Fokus auf der physiologischen Meß- und Bestimmbarkeit der auditorischen Leistung.

Segmentierung anhand von Silbengrenzen

Ein Beispiel aus dem Englischen belegt, wie wichtig die Detektion von Silbengrenzen für die Interpretation eines Wortes ist. Man vergleiche hierbei die unterschiedliche Silbifizierung in:

(1) *an aim* vs. *a name*

In dieser Sequenz führt nur das unterschiedliche Pausenmuster zu einem Bedeutungsunterschied.

Segmentierung anhand von Prominenzmerkmalen

Auch Prominenzmerkmale können zur Strukturierung der eintreffenden Signale akustischer Ereignisse genutzt werden. Hierbei kann allein eine Betonungsverschiebung zur unterschiedlichen Bedeutung eines Wortes beitragen, wie folgendes Beispiel demonstrieren soll:

(2) *UmFAHren* vs. *UMfahren*.

Die eben genannten Beispiele sollen verdeutlichen, wie relevant das Erkennen prosodischer Information unterschiedlichen Typs für die Interpretation von Sprache sein kann. Diese prosodische Information wird hauptsächlich durch die vielfältige Variation segmentaler und suprasegmentaler akustischer Parameter übertragen. Dabei ist nicht genau klar, an welchen akustischen

Merkmale sich das auditorische System orientiert, um eine Segmentierung und damit Identifizierung zu bewerkstelligen, da nicht nur ein Parameter separat verändert wird, sondern mehrere kumuliert. Die Betonung der Silbe, wie das /UM/ in UMfahren, wird produziert, indem sich die Dauer, Intensität und Grundfrequenz der ersten Silbe in Beziehung zu den restlichen Silben in ihren Merkmalen verändern bzw. erhöhen (prominenter werden). Anhand einer elektrophysiologischen Studie sollen Belege für neuronale Reaktionen während der Verarbeitung von Domänengrenzen und Prominenzmerkmale untersucht werden.

Stimulusmaterial

Die Bedingungen, die sich hierfür an das akustisch darzubietende Material stellten waren, dass es:

- natürlich-sprachliches Material sein sollte,
- Silbengrenzen vorhanden und auch akustisch nachweisbar sind und
- Betonungen an unterschiedliche Positionen auftreten können.

Realisiert wurde dies durch die Präsentation einer homogenen Konsonant-Vokal-Struktur aus drei Silben mit einem akustisch gut segmentierbaren Konsonanten als Onset und einem Vollvokal. Dazu eigneten sich die alveolaren Plosive /d/ und /t/ und der Vokal /a/. Eingesprochen wurden die Dreisilber von einer weiblichen Sprecherin wobei die Betonungsmuster von deutschen Nomen als Vorlage benutzt und nachgesprochen wurden (Erstbetont: Kolibri, Zweitbetont: Minute). In Bild 1 wird ersichtlich, dass die Betonung der initialen bzw. medialen Silbe vor allem eine Dehnung des Vokals zur Folge hat. Hier wird auch ersichtlich, dass sich das natürliche Betonungsmuster nicht an einem metrischen Konzept orientiert. Die finale Silbe ist jeweils die längste. Daraus alleine läßt sich schließen, dass die Prominenz der vorangegangenen Silbe durch andere Parameter wie Frequenz und Intensität zusätzlich ausgedrückt werden muß.

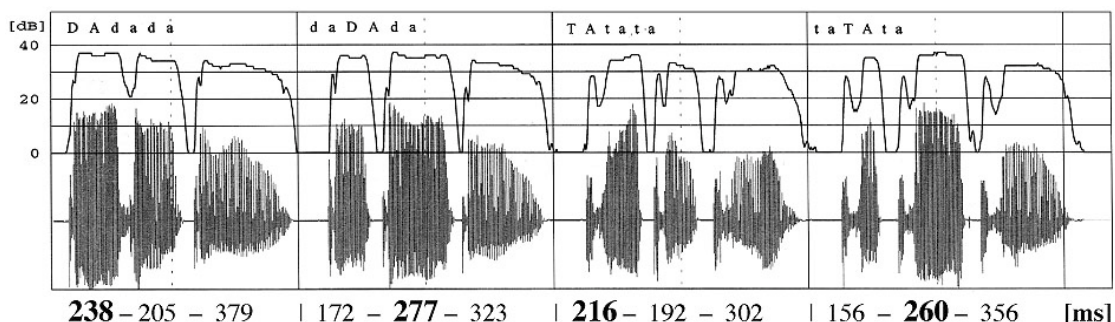


Bild 1: Oszillogramm und Amplitudenhülle des Stimulusmaterials bestehend aus zwei verschiedenen Dreisilbern mit initialer und medialer Betonung. Die Dauer der einzelnen Silben ist in ms angegeben. Die betonte Silbe (fett markiert) ist vor allem durch die Längung des Vokals gekennzeichnet. Die finale Silbe bleibt bei der natürlichen Aussprache trotz Prominenz auf der ersten oder zweiten Silbe die längste Silbe.

Method

Mit Hilfe der zeitlich hoch auflösenden, nicht invasiven Messung des Elektroenzephalogramms (EEG) kann die neuronale Segmentierungsleistung des auditorischen Inputs sichtbar gemacht werden. Die Darbietung sich wiederholender akustischer Stimuli (Standard) und einem darauffolgenden abweichenden Reiz (Deviant) löst in ein bestimmtes Ereigniskorreliertes Potential (EKP) aus, die Mismatch Negativity (MMN) [1]. Da diese Komponente in einem frühen Zeitfenster von 150-200 ms nach Stimulusonset auftritt und unabhängig von der Aufmerksamkeit ist, geht man davon aus, dass in dieser Komponente ein exogener Prozess reflektiert wird, der allein auf der neuronalen Verarbeitung des akustischen Inputs beruht.

Bei der Darbietung wurde jeder Stimulus als Standard und als Deviant eingesetzt. Um die Probanden von der Devianterkennung abzulenken, wurde ihnen die Aufgabe gestellt, die Stimuli blockweise zu zählen. Dabei wurden jeweils 8-10 aufeinanderfolgende Stimuli präsentiert.

Neuronale Verarbeitung von Silbengrenzen

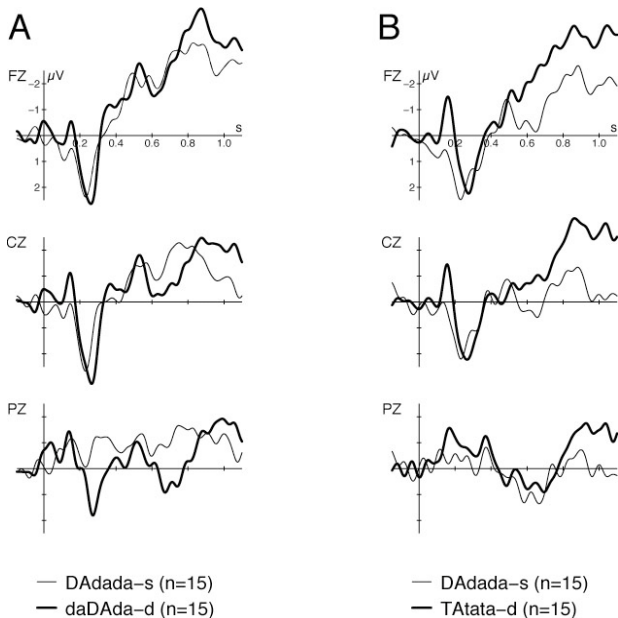


Bild 2: Mittelung der EKPs von 15 Probanden als Reaktion auf den A: Metrikwechsel und B: Phonemwechsel. Standard (grau), Deviant (schwarz).

In Bild 2 sind zwei der 12 untersuchten Bedingungen dargestellt. Sie zeigen, dass im Zeitfenster von 150-200 ms eine Negativierung (nach oben abgetragen) nach dem Stimulusonset auftritt. Diese ist bei der Darbietung des Betonungswechsels (Metrikwechsel) für die deviante Bedingung ausgeprägter als für den Standard und löst beim Phonemwechsel die stärkste Reaktion hervor. Die beim Metrikwechsel ausgelöste parietale Positivierung nach 250 ms tritt bei dem Phonemwechsel ohne Metrikwechsel nicht auf.

Der Vergleich der erst- und zweitbetonten Stimuli in Bild 3 zeigt, dass die Silbengrenzen bzw. -onsets auch in späteren Zeitfenstern zusätzliche Negativierungen auslösen. Dabei ist diese in der devianten Bedingung stärker ausgeprägt. Was darauf hinweist, dass die präsentierten Stimuli differenziert werden.

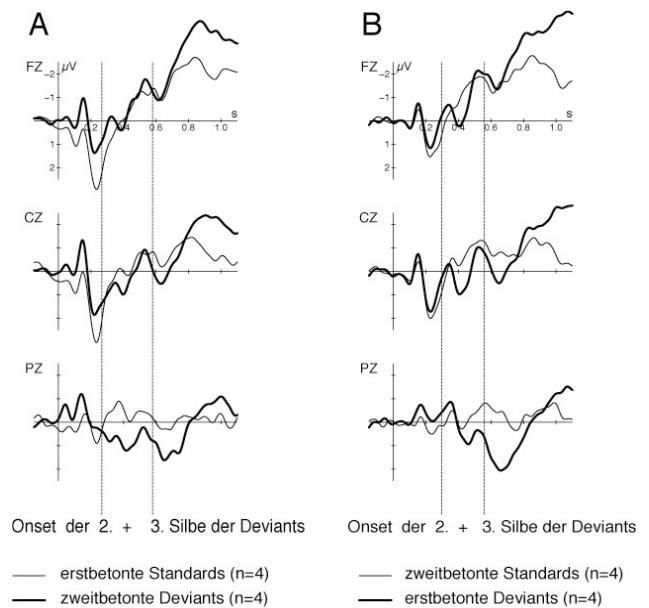


Bild 3: Mittelung aller EKPs als neuronale Antwort auf die erstbetonten Stimuli (A: Standard, B: Deviant) im Vergleich zu den zweitbetonten Stimuli (A: Deviant, B: Standard). Die vertikalen Linien zeigen die gemittelten Onsets der 2. und 3. Silbe der Deviants.

Aussicht

Psychoakustische Untersuchungen zeigen, dass trotz starker Manipulation der spektralen Eigenschaften des sprachlichen Materials (Reduzierung der Bandbreiten) [2] eine Worterkennung immer noch möglich ist. Auch die Fähigkeit, mit fehlender Frequenzinformation bei einem Cochleaimplantat Sprache zu verstehen [3], weist darauf hin, dass die Verarbeitung der Amplitudenhülle eine Konstante innerhalb des Hörvermögens darstellt und die komplexe Frequenzinformation dabei eine unterstützende Rolle spielt [4]. In weiteren Experimenten soll methodisch die Information der Amplitudenhülle extrahiert werden und die neuronale Antwort im EEG untersucht werden [5, 6].

[1] Näätänen, R. (1990): The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *The Behavioral and Brain Science*, Vol. **13**, 201-288.

[2] Shannon, R.V., Zeng, F., Kamath, V., Wyngonski, J. and Ekelid, M. (1995): Speech recognition with primarily temporal cues. *Science Reports*, Vol. **270**, 303-304.

[3] Kraus, N., Micco, A.G., Koch, D.B., McGee, T. Carrell, T., Sharma, A., Wiet, R.J. and Weingarten, C.Z. (1993): The mismatch negativity cortical evoked potential elicited by speech in cochlear-implant users. *Hearing Research*, **65** (1-2): 118-124.

[4] Tervaniemi, M., Schröger, E., Saher, M. and Näätänen, R. (2000). Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans – a mismatch negativity study. *Neuroscience Letters*, **290**: 66-70.

[5] Hies, T. (2001): *Akustisch evozierte Hirnrindspotenziale: Sprachähnliche Stimulation und Artefaktkorrektur*. Diss. Technische Fakultät, Universität Erlangen-Nürnberg, Shaker Verlag.

[6] Kruck, S., Berti, S. and Alter, K (2001): *Speech specific auditory evoked potentials*. Fechner Day 2001, Pabst Science Publishers.