

DAUERPHÄNOMENE IN LISTENFÖRMIGEN AUFZÄHLUNGEN

Hannes Pirker

Österreichisches Forschungsinstitut für Artificial Intelligence (ÖFAI) *
Schottengasse 3, A-1010 Wien
Email: hannes@ai.univie.ac.at

1 EINFÜHRUNG

Diese Arbeit stellt eine korpusbasierte Untersuchung von listenförmigen Aufzählungen in Wetterberichten vor. Dabei werden primär Phänomene der Dauervariation betrachtet.

Die Wahl dieses Äußerungstyps für den Aufbau eines Korpus wurde durch mehrere Überlegungen motiviert: Aufzählungen stellen einen sehr spezialisierten Äußerungstyp dar, dessen Modellierung insbesondere für sprachbasierte Auskunftssystemen interessant ist. Die Korpusanalyse soll als Grundlage für eine möglichst natürliche Synthese solcher Äußerungen innerhalb eines Concept-to-Speech Systems dienen. Diese Textsorte weist eine weitgehend festgelegte Struktur und einen relativ geringen Wortschatz auf, sodaß eine gute Vergleichbarkeit der Daten gegeben ist. Da diese Aufzählungen zudem mehrmals täglich im Rundfunk übertragen werden, bietet sich die Gelegenheit mit vertretbarem Aufwand einen Korpus mit sehr ähnlichen Inhalten von einer größeren Anzahl von Nachrichtensprechern aufzubauen.

2 WETTERLISTEN

Die ausgewerteten Meldungen weisen folgende Standardstruktur auf: Eine Meldung besteht aus „Meldungszeilen“, je eine Zeile pro österreichischer Landeshauptstadt. Ein Beispiel:

- (z1) *Wien heiter 13 Grad,*
- (z2) *Eisenstadt heiter 14,*
- (z3) *St.Pölten und Linz heiter 16,*
- (z4) *Salzburg stark bewölkt 16,*
- ...
- (z5) *und Klagenfurt heiter 12 Grad.*

Die typische Intonationsstruktur solcher Meldungen läßt sich nach G-ToBI ([Reyelt et al. 96]) folgendermaßen beschreiben: Jede Zeile bildet eine Intonationsphrase, die mit einem hohen Grenzton (H%) endet. Der Städtenamen und die Ziffern der Temperaturangabe sind mit einem Gipfelakzent (H*) versehen, während die Bewölkungsangabe üblicherweise unbetont bleibt.

* Das ÖFAI wird vom *Österreichischen Bundesministerium für Wissenschaft und Verkehr* unterstützt. Die gegenständliche Arbeit wurde zudem vom *Fonds zur Förderung der wissenschaftlichen Forschung (FWF)* im Rahmen des Projekts P10822 gefördert.

3 DATENKORPUS UND AUSWERTUNGSTECHNIK

Der Korpus besteht zur Zeit aus Aufnahmen von 21 Sprechern und umfaßt 72 Meldungen mit einer Gesamtlänge von 22 Minuten. Es wurden mittels *Speech Filing System SFS*¹ manuell Wortmarken vergeben und die Wortgrenzen markiert. Außerdem wurden Pausen, die auch ein Phrasenende markieren, mit /b/, alle anderen mit /p/ markiert. Der Korpus umfaßt knapp über 4000 annotierte Einheiten.

Um die starken Schwankungen der Sprechrate etwas auszugleichen, werden die ermittelten Dauerwerte normalisiert: Für jede Meldung wurde die durchschnittliche Länge der Städtenamen ermittelt und alle Dauern als Prozentsatz dieses Wertes angegeben. Sollen Dauern aus unterschiedlichen Meldungen miteinander in Beziehung gesetzt werden, wird also auf diese normalisierten Werte und nicht auf absolute Längen in [ms] zurückgegriffen.

Die Labeldaten sind unter Prolog in einer einfachen relationalen Datenbasis repräsentiert (vgl. etwa [Draxler 95]). So werden auch zusätzliche Informationen wie Zugehörigkeit zu semantischen und syntaktischen Klassen, Silbenstruktur etc. zugeordnet.

Zur schnellen Überprüfung verschiedener Hypothesen wurde ein Programm erstellt, das einfache statistische Auswertungen vornimmt und insbesondere Histogramme wie in Abb.2 oder Abb.3 erzeugt. Diese geben nicht nur die Längen der einzelnen Wörter graphisch wieder, was einen schnellen Überblick über Dauerverteilungen ermöglicht, sondern machen bei „Anklicken“ eines Balkens auch die dazugehörigen Audiodaten hörbar. Das ermöglicht die sofortige perzeptive Evaluation von Wörtern, die in der grafischen Längendarstellung Aufmerksamkeit erregen.

4 AUSWERTUNGSBEISPIELE

Ein besonderes Augenmerk wurde auf den Einfluß der Phrasierung auf die Dauer gelegt.

Prefinale Längung

Das Phänomen der prefinalen Längung vor Phrasengrenzen ist allgemein bekannt. Es wurde untersucht, inwiefern sich dieser Effekt in der sehr einfachen Repräsentation unseres Korpus nachweisen läßt. Da nur die Wortgrenzen zur Verfügung stehen, besteht ja z.B.

¹<http://www.phon.ucl.ac.uk:80/resource/sfs.html>

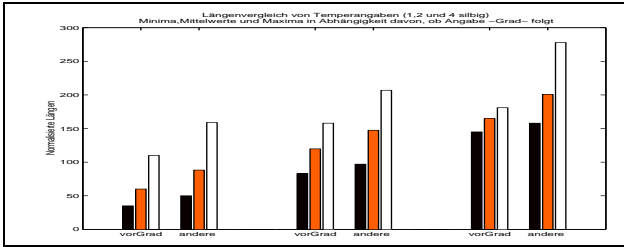


Abbildung 1: Längenvergleiche von Temperaturangaben in Abhängigkeit davon, ob sie von /grad/ gefolgt werden oder in phrasenfinaler Position stehen. Die drei Gruppen beziehen sich auf ein-, zwei- sowie viersilbige Zahlen. (grau = Mittelwert, schwarz= Minimum, weiß = Maximum).

keine Möglichkeit den Einfluß auf verschiedene Silben direkt zu evaluieren. Zudem ist die Markierung der Phrasen sehr rudimentär: lediglich Phrasengrenzen, die auch eine Pause aufweisen, sind explizit mit einem /b/ markiert.

Das Wort /grad/ findet sich immer nur in unbetonter, phrasenfinaler Position (z1,z5) und bietet sich daher für eine Untersuchung der prefinalen Längung an. Die statistische Auswertung ergab dabei, daß sich *kein* Einfluß des nachfolgenden Labels auf die Dauer nachweisen ließ: die durchschnittlichen Dauern blieben konstant, egal ob danach eine explizite Pause /b/, ein /end/ (markiert das Ende einer Gesamtmeldung wie (z5)) oder ein pausenloser Übergang zur nächste Zeile erscheint. An dieser speziellen, auch syntaktisch und semantisch eindeutigen Phrasengrenze hat die An- oder Abwesenheit einer expliziten Pause also keinen Einfluß auf die Länge des vorhergehenden Wortes.

Dieses Ergebnis wurde auch bei der Auswertung präfinaler Temperaturangaben (wie in (z2,z3,z4)) bestätigt.

Der Einfluß der präfinalen Längung ist insgesamt jedoch sehr deutlich feststellbar. In Abb.1 zeigt sich, wie sich die mittleren Längen der Temperaturangaben unabhängig von ihrer Silbenanzahl um einen relativ konstanten Wert erhöhen, wenn sie direkt am Phrasenende stehen, und nicht von /grad/ auf „Abstand“ gehalten werden.

Phrasenlänge

Die teilweise sehr rhythmisch wirkende Artikulation von Listaufzählungen legt es nahe, Isochronieeffekte zu untersuchen. Hierzu wurden koordinierte Konstruktionen wie in (z3) analysiert: Unterscheiden sich die Angaben zu adjazenten Städten nicht, so werden diese üblicherweise zusammengefaßt. Die These war, daß die Städtenamen in solchen Konstruktionen gekürzt realisiert werden.

Abb.2 zeigt die Dauerverteilungen einiger Städte grafisch. Die statistische Auswertung zeigt einen eindeutigen Kürzungseffekt bei denjenigen Städtenamen, die an erster Stelle (also *vor* dem /und/) stehen. Der direkte Vergleich einiger solcher Daten in Abb.3 zeigt wiederum, daß die Kürzungen im zweiten Teil der NP weitaus geringer ausfallen: hier findet vermutlich eine Überlagerung mit der Tendenz statt, nach solchen

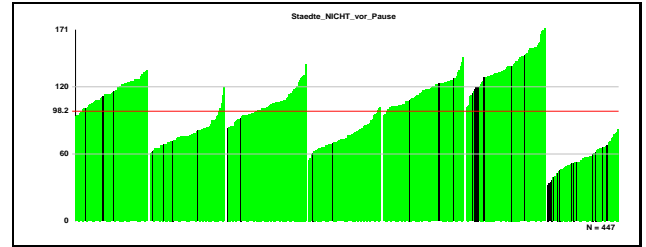


Abbildung 2: Städtenamen, die nicht von einem /b/ gefolgt werden, nach Namen und normalisierter Länge sortiert. (schwarz = stehen vor /und/, grau = sonstige)

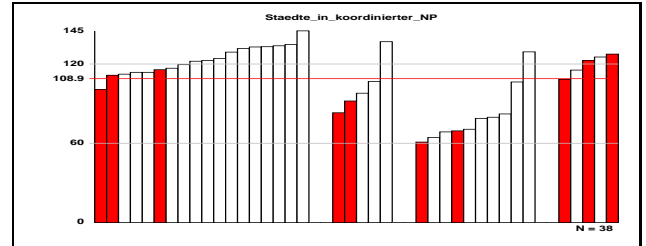


Abbildung 3: Auswahl von Städtenamen in koordinierten Nominalphrasen. (grau = an 1. Stelle, weiß = an 2. Stelle)

NPs eine zusätzliche prosodische Grenze einzufügen, und damit präfinale Längung auszulösen.

5 DISKUSSION UND AUSBLICK

Die gegenständliche Arbeit wurde u.A. in Hinblick auf den geplanten Aufbau eines größeren Korpus zur Untersuchung von Segmentlängen durchgeführt, um im kleinen Maßstab die Eignung verschiedener Repräsentationsformen zu evaluieren. Insbesondere die Implementierung der Datenbank in Prolog hat sich bestens bewährt. Für die weitere Arbeit ergab sich, daß insbesondere auf die Repräsentation prosodischer Grenzen größeres Augenmerk zu legen ist. In diesem Zusammenhang soll in einem nächsten Schritt überprüft werden, wie dies am Besten im Zusammenspiel mit einer phonologischen (ToBI) oder parametrischen (z.B. [Portele et al. 95]) Repräsentation der Grundfrequenzverläufe realisiert werden kann. Die angewendeten einfachen statistischen Methoden haben sich für die Durchführung der vorgestellten Untersuchungen noch als praktikabel erwiesen, für multifaktorielle Auswertungen werden jedoch in Hinkunft ausgefeiltere Verfahren auszuwählen sein.

Literatur

[Draxler 95] Draxler C.: Introduction to the Verbmobil-PhonDat Database of Spoken German, Practical Applications of Prolog Conference 95, Paris, 1995.

[Portele et al. 95] Portele T., Krämer J., Heuft B., Sonntag G.: Parametrisierung von Grundfrequenzkonturen, in Fortschritte der Akustik, DAGA-95, 1995.

[Reyelt et al. 96] Reyelt M., Grice M., Benz Müller R., Mayer J., Batliner A.: Prosodische Etikettierung des Deutschen mit ToBI, in Gibbon D.(ed.), Natural Language Processing and Speech Technology, Mouton de Gruyter, Berlin, p.144-155, 1996.