

# **Phonsegmentierte, nichtlineare Zeitskalierung von Sprache**

## *Algorithmen und Anwendung auf Wahrnehmungsuntersuchungen sehbehinderter Probanden*

Die Arbeit beschäftigt sich mit der zeitlich skalierten Sprache, insbesondere deren Beschleunigung. Akzelerierte Sprachausgaben stellen insbesondere für sehbehinderte Menschen ein wichtiges Kommunikationsmittel dar. Zunächst wird ein Überblick über die Terminologie und Definitionen bezüglich der Sprechgeschwindigkeit gegeben. Ausgehend vom allgemeinen Kommunikationsbegriff werden neben der Sprachkommunikation insbesondere die temporalen Limitierungen der Sprachproduktion und Sprachperzeption erörtert. Hier wird die natürliche Be- und Entschleunigung der Sprechgeschwindigkeit betrachtet. Auch die Wahrnehmung zeitlich skalierten Sprache, ob natürlich oder apparativ, ist limitiert. Die Grenzen der Wahrnehmbarkeit beschleunigter akustischer Sprache hängen unter anderem auch von der Sehfähigkeit des Hörers ab. Aus diesem Grund werden insbesondere bereits veröffentlichte und eigene Versuchsergebnisse mit sehbehinderten, blinden und sehenden Menschen präsentiert. Auch die Nutzung der skalierten Sprache in vielen akustischen Kommunikationshilfsmitteln der blinden Nutzer spielt dabei eine Rolle.

Der Kernpunkt dieser Arbeit beschäftigt sich mit der instrumentellen Skalierung von Sprache. Neben dem historischen Rückblick werden aktuelle Skalierungsmethoden und -algorithmen betrachtet und eine mögliche Strategie zur Verbesserung bezüglich der Verstehbarkeit und Verständlichkeit erarbeitet. Auf Grund dieser Strategien wird ein neuer nichtlinearer Skalierungsalgorithmus entwickelt. Die Analyse, Segmentierung und Dezimierung des Sprachsignals arbeitet zur Erhaltung der Echtzeitfähigkeit ausschließlich in der zeitlichen Domäne. Die neuen Sprachskalierungsstrategie beinhalten die Erhaltung aller Lautübergänge und führt zur ausschließlichen Reduktion der Phonnukleus. Die relative Menge der eliminierten Signalanteile hängt von der Detektion einer stimmhaften, stimmlosen oder fehlenden Anregung ab.

Mit den auditiven Experimenten werden zunächst mit Hilfe eines bereits bekannten linearen Skalierungsverfahrens die Wahrnehmungsfähigkeiten von sehenden, sehbehinderten und blinden Probanden bezüglich Verstehbarkeit und Verständlichkeit untersucht. Weitere instrumentelle und auditive Analysen dienen der Parameterevaluierung des in dieser Arbeit entwickelten nichtlinearen Skalierungsverfahrens. Abschließend erfolgt die auditive Evaluierung des phonsegmentierenden, nichtlinearen Sprachskalierers mit aktuellen Skalierungsverfahren. Weiterhin wird auch neben der zeitlichen Skalierung die Grundfrequenz in ihrer Dynamik bis zur Monotonie eingeschränkt, um die Frage zu untersuchen, ob die Verringerung der Grundfrequenzdynamik eines beschleunigten Sprachsignals dessen Verstehbarkeit und Verständlichkeit erhöht.

# **Phone Segmented Nonlinear Time Scaling of Speech**

## *Algorithms and Utilization in Perceptual Tests with Visually Impaired Persons*

This PhD thesis deals with time scaled speech, particularly with accelerated speech. Accelerated speech is an important feature of the communication for people living with visual impairments or blindness.

In the first chapter, an overview of the terminology and definition concerning accelerated speech will be defined. Beginning with the common concept of the communication, the temporal limits of speech production and speech perception will be discussed along with considerations of the natural increase or decrease of the speech rate.

The second chapter lists the possibilities and limitations of accelerated speech perception and speech production. The limits of the perception of natural or instrumental accelerated speech are independent from the visual capabilities of the listeners. For this reason the results of published experiments and own experiments from this group of people are presented. These results will be compared with the auditive perception of sighted people in order to see how the utilization of accelerated speech plays an important role in communication aids for blind people.

The core issue of this thesis deals with the instrumental acceleration of speech signals. Besides the historical retrospect, the thesis considers the current algorithms for time scaled speech and develops a strategy to improve the comprehension and intelligibility of accelerated speech. Based on this strategy a new nonlinear algorithm to scale the speech rate will be developed. The analysis, segmentation and classification of the algorithm work only in the time domain to obtain the real time capability. The strategies of scaling will be used for the retention of the phone transition and to exclusively reduce the speech signal in the phone nuclide only. The relative amount of eliminated parts of speech signals depends on voiced, voiceless or none excitations.

In the subjective hearing experiments, the perception abilities of sighted, visually impaired or blind people will be analyzed with a known linear speech time scaler. Further objective and subjective analysis compares the parameters of this developed nonlinear algorithm and also evaluates this new algorithm with the actual linear algorithm working in the time domain exclusively. Furthermore the pitch dynamic in accelerated speech is decreased to examine the question if manipulation of pitch will increase the intelligibility and comprehension whenever the speech is accelerated as well.