



Agentensimulationen und KNNs in der Linguistik

VON MARTIN PYKA

Inhaltsverzeichnis

I	Einleitung.....	3
II	Talking Heads.....	5
III	Experimente mit künstlichen neuronalen Netzen.....	8
III.1	Sentence Processing mit KNNs	8
III.2	Sequentielle Satzverarbeitung mit KNNs	9
IV	Ausblick.....	11

I Einleitung

Nachdem über 150 Jahre lang die Sprachwissenschaft weitgehend autonom versuchte, sprachliche Phänomene zu erklären und in ihrer Gesamtheit einzuordnen, entwickelt sich diese Aufgabe seit etwa zwei Jahrzehnten zu einem interdisziplinären Forschungsgebiet. Und das ist auch dringend notwendig. Eine Vielzahl von Branchen haben ein grosses Interesse daran, Kommunikationssysteme auf Basis natürlicher Sprache verfügbar zu machen, um dem Menschen beispielsweise im Auto, an Serviceschaltern oder zu Unterhaltungszwecken die Möglichkeit zu bieten, durch natürliche Sprechakte Kontrolle auszuüben. Neben den sogenannten Dialogsystemen sind es vor allen Dingen Übersetzungsprogramme, an denen am intensivsten gearbeitet wird, um eine länder- und völkerübergreifende Verständigung zu ermöglichen.

Dahinter stehen zum Teil auch ganz konkrete politische Interessen. Im EU-Parlament werden alle Debatten in die elf Amts-Sprachen des EU-Parlamentes übersetzt. Alle beschlossenen EU-Gesetze, Richtlinien, internationale Übereinkommen, Pressemitteilungen, technische Studien, Finanzberichte, Sitzungsprotokolle und vieles mehr müssen sogar in alle Sprachen der Mitgliedsländer übersetzt werden. Bei der manuellen Übersetzung gibt es eine Vielzahl von Faktoren, die den Übersetzungsvorgang erschweren oder stören. Interpretationsspielräume in der Ausgangssprache, Probleme bei der direkten Übersetzung in die Zielsprache, Ambiguitäten in der Zielsprache, all dies kann zu einer inhaltlichen Verzerrung der ursprünglichen Texte führen, schliesslich liegt es in der Hand des Übersetzers, den Originaltext in die Zielsprache zu transferieren. Nicht immer kommen also zwei Übersetzer, die den gleichen Text übersetzen sollen, zu der gleichen Übersetzung. Wünschenswert wäre es daher, wenn genau so eine 1:1-Abbildung von Gesetzestexten automatisiert herstellbar wäre. Die inhaltliche Auseinandersetzung könnte auf diese Weise ohne sprachliche Barrieren und Reibungsverlusten durch manuelle Übersetzungen stattfinden. Desweiteren würde eine automatisierte Übersetzung die Kommunikationswege enorm beschleunigen und nicht zuletzt teures Personal sparen.

Es gibt daher eine ganze Reihe Projekte, die sich dieser Problematik annehmen und nach Systemen forschen, die eine adäquate Übersetzung, entweder innerhalb einer bestimmten sprachlichen Domäne oder allgemein, zulassen. Zu den nennenswertesten Vertretern unter den Übersetzungsprogrammen zählen sicherlich Verbmobil und Systran. Verbmobil war ein langfristig angelegtes Leitprojekt des Bundesministeriums für Bildung und Forschung und sollte die Kommunikation zwischen den Sprachen Deutsch, Englisch und Japanisch in den Domänen Termin-, Reiseplanung und Hotelreservierung vereinfachen. Es wurde von 1993 bis 2000 entwickelt und gilt auch heute noch als State of the Art in diesem Bereich. Systran ist hingegen ein kommerzielles Produkt, das als universelle Übersetzungshilfe für den alltäglichen Gebrauch Verwendung finden sollte. Es lassen sich für über 100 Sprachen Übersetzungspakete dazu kaufen.

Beiden Systemen ist gemeinsam, dass sie Sprache weitgehend auf Symbolebene verarbeiten und nur bedingt lernfähig sind. Allen vorhandenen Übersetzungssystemen ist gemeinsam, dass sie nach wie vor nicht zufriedenstellende Ergebnisse liefern. Zu leicht lassen sich Wendungen und Formulierungen finden, die das System aus der Bahn werfen und nicht nur zu sinnentstellten, sondern zum Teil auch zu ungrammatischen Übersetzungen führen. Es stellt sich also die Frage, inwieweit sich mit wissensbasierten Systemen, die Sprache auf Symbolebene verarbeiten, tatsächlich eine Anwendung erstellen lässt, die in der Lage ist, Übersetzungen anzufertigen, die vom Weltwissen und den Erfahrungen von Menschen zeugen.

In den folgenden Kapiteln sollen deshalb Wege vorgestellt werden, wie sich mit Hilfe von Agentensystemen und neuronalen Netzen Sprache bzw. sprachliche Phänomene simulieren lassen. Im Vordergrund steht dabei die Frage, wie eigentlich der Mensch Sprache erlernt und verarbeitet und wie sich diese Mechanismen am Computer simulieren lassen. Noch ist dieser Forschungsbereich sehr jung und klein und brachte noch keine praktischen Anwendung, geschweige denn, nennenswerte Übersetzungssysteme hervor. Die Untersuchung der menschlichen Sprachverarbeitung offenbart aber schon jetzt, dass sprachverarbeitende Software mehr berücksichtigen muss, als die verschriftlichten oder aufgezeichneten Sprechakte, mit denen der Linguist bisher gearbeitet hat.

II Talking Heads

Die Talking Heads ist eine von Luc Steel¹ entwickelte Agentensimulation, bei dem durch die Lernregeln jedes einzelnen Agenten ein kollektiver Wortschatz entsteht. Jeweils zwei Agenten schauen über eine Videokamera auf ein Whiteboard, auf dem geometrische Objekte angebracht sind. Diese unterscheiden sich nach Lage, Farbe, Grösse und Form. Die Agenten sind in der Lage das Bild nach diesen Kategorien zu segmentieren und Konzepte, wie „oben“, „links“, „eckig“ herauszubilden. Ein Agent dieses Naming-Games, in das zwei Agenten involviert sind, nimmt die Rolle des Sprechers ein, der andere die Rolle des Hörers. Der Sprecher benennt nun ein Objekt eindeutig, in dem er, sofern noch nicht vorhanden, die nötigen Informationen, wie zum Beispiel „ganz links – unten – blau“, durch ein Phantasiewort kodiert und dieses dem anderen mitteilt. Der Hörer muss nun versuchen zu erraten, welches Objekt der Sprecher gemeint haben könnte. Zeigt er auf das korrekte Objekt, verstärken beide Agenten das gewählte Wort für die betroffenen Kategorien. Der Hörer-Agent besetzt dabei alle Kategorien, in denen das gewählte Objekt vorhanden ist, mit dem neu erlernten Wort. Findet der Hörer das korrekte Objekt nicht, zeigt ihm der Sprecher, welches Objekt er gemeint hat. Im Wortschatz beider Agenten findet dann ein Korrekturprozess statt, der beim Sprecheragenten die Wertigkeit des Wortes für die gewählte Kategorien abstuft und beim anderen erhöht.

Durch viele weitere Wortspiele, in denen die Agenten jeweils in die Sprecher und Hörerrolle schlüpfen und mit anderen Agenten in Kontakt kommen, breiten sich neu erfundene Wörter von Agent zu Agent aus, aber verändern sich auch dabei. Untersucht man nach tausenden von Sprachspielen, welche Wörter jeder einzelne Agent kennt bzw. wie sich die Popularität eines Wortes im Verlauf der Simulation verändert hat, lassen sich Entwicklungen feststellen, die sich auch in realen Sprachen finden.

So ist es grade am Anfang einer Simulation verständlich, dass zunächst mehrere Begriffe für das gleiche Konzept entstehen, die sich über die gesamte Population ausbreiten. Der prozentuale Anteil dieser Synonyme verändert sich, je nachdem welche Konstellationen bei den Sprachspielen entstehen. In Abbildung II.1 kann man erkennen, wie zunächst eine ganze Reihe von Begriffen einer Bedeutung zugeordnet ist. Zunächst ist das Wort *danwe*

¹ Luc Steel ist Professor of Computer Science an der freien Universität in Brüssel. Zu seinen wissenschaftlichem Interessengebiet zählen alle Arten von Künstlicher Intelligenz, unter anderem Natural Language, Vision, Robot Behaviour, Learning, Cognitive Architecture und Knowledge Representation; <http://arti.vub.ac.be/~steels/>

dominant. Durch eine günstige Kombination einiger Sprachspiele gewinnt jedoch irgendwann das Wort *wogglespiat* eine derart grosse Popularität, dass alle weiteren Wörter nahezu verdrängt werden.

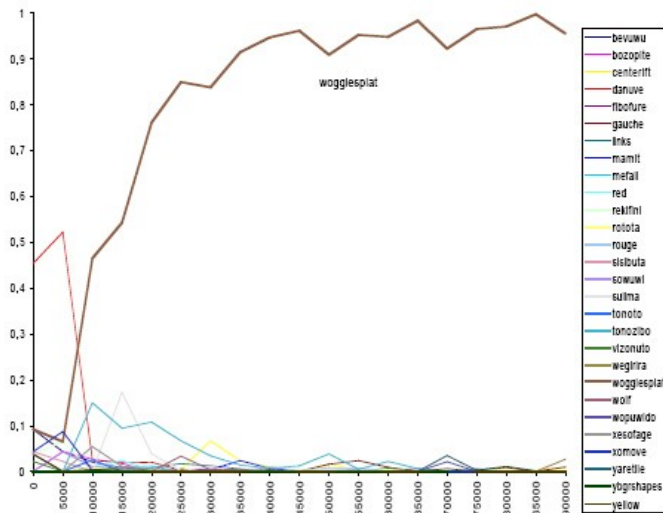


Abbildung II.1: Prozentualer Anteil der Wörter, die die gleiche Bedeutung haben

Auch Polysemie (ein Wort hat mehrere Bedeutungen) liess sich feststellen und leicht erklären. Bei Agenten, die ein neues Wort mit den falschen Objekten in Verbindung bringen, kann sich unter Umständen eine andere Bedeutung heraus kristallisieren, die dann in einer Teilgruppe der Population weitere Verbreitung findet. Durch Wortspiele zwischen den Agenten beider Gruppen vermischen sich auch die Wortbedeutungen. Insbesondere für genauere Lagebeschreibungen, wie zum Beispiel „Mitte – Links“ und „Ganz links“ wurde öfters das gleiche Wort benutzt, bevor es sich nach einer bestimmten Anzahl an Wortspielen auf eine Bedeutung beschränkte.

Luc Steels Talking Heads sind ein anschauliches Beispiel für eine neuere linguistische Theorie, die unter dem Namen „Evolutionary Linguistics“ bekannt ist. Sprache ist in dieser Theorie das emergente Produkt kommunikativer Grundregeln. Jeder Agent ist dabei stets darum bemüht, die Kommunikation mit seinen Nachbarn so sicher und leicht wie möglich zu halten. Durch jeden getätigten und wahrgenommenen Sprechakt passt er sein eigenes Sprachwissen an das der anderen an. Durch „kreative Fehler“ und Innovationen verändert er aber das Sprachwissen der anderen aktiv mit. Sprache ist also kollektives Wissen, das sich durch Konventionalisierung und evolutionäre Weiterentwicklung ständig verändert. Diese Theorie schliesst andere Aspekte des menschlichen Lebens ausdrücklich mit ein. Insbesondere die kognitiven und motorischen Fähigkeiten werden als grundlegende Voraus-

setzungen gesehen, um komplexes Verhalten und dadurch auch Sprache entstehen lassen zu können. Die evolutionäre Linguistik kann daher als interdisziplinärer Forschungsbereich verstanden werden, der vor allen Dingen Erkenntnisse aus anderen wissenschaftlichen Bereichen, wie der Verhaltensforschung, der Psychologie oder der Biologie, zusammenbringen will.

III Experimente mit künstlichen neuronalen Netzen

Agentensimulationen, in denen die Komplexität erst über die gesamte Gruppe entsteht, jeder Agent aber ein Regelsystem für das Verhalten besitzt, werden wohl kaum ausreichen, um alle Facetten von Sprache untersuchen zu können. Daher erscheint es reizvoll, die mustererkennenden und lernenden Fähigkeiten des Menschen am Computer zu simulieren. In diesem Zusammenhang haben sich künstliche neuronale Netze als abstrahiertes Modell der realen Hirnstruktur für derartige Problemstellungen etabliert. Im folgenden werden zwei konkrete Modelle vorgestellt, mit denen es gelang, Sprachwissen, wenn auch in einem sehr eingeschränkten Maß, durch Lernen entstehen zu lassen.

III.1 Sentence Processing mit KNNs

Das 1996 von McClelland und Kawamoto entwickelte künstliche neuronale Netz ist in der Lage durch überwachtes Lernen die Konstituenten eines Satzes thematischen Rollen zuzuordnen. So ist beispielsweise *the boy* in dem Satz *The boy broke the window* ein Agens, während *the stone* in *The stone broke the window* ein Instrument ist. Das Netz wurde also mit Sätzen trainiert, die eine einfache Oberflächenstruktur aufweisen, beispielsweise Subjekt – Prädikat, Subjekt – Prädikat – Objekt, oder Subjekt – Prädikat – Objekt – Präpositionalphrase. Das Netz bekam dabei die Wörter als sogenannte Merkmalsvektoren übergeben, in denen Nomen nach Kategorien, wie *human*, *softness* oder *gender* kodiert waren und Verben nach Kategorien, wie *doer*, *cause* oder *touch*. Ein Microfeature, also ein konkretes Merkmal, konnte die Werte 1, 0 oder ein Fragezeichen haben, wenn die Kategorie irrelevant ist oder eine Ambiguität vorliegt. Als Ausgabe sollte das Netz die Kasus-Struktur des Satzes zurückliefern. Bei einem Satz wie *The boy broke the window with the hammer* wäre demnach *the boy* der Agens, *the window* der Patient und *the hammer* das Instrument.

Das Netz wurde mittels der Delta-Regel trainiert. Auf Grund der Tatsache, dass keine bloße Zeichenkette dem Netz übergeben wurde, sondern semantische Informationen, gelang es dem Netz nach nur 50 Trainingszyklen auch bei neuartigen Sätzen mit einer Zuverlässigkeit von 98,5% vorrauszusagen, welche Rolle die Konstituenten des Satzes einnehmen. Die Semantik des Satzes fand dabei mehr Berücksichtigung, als die Satzstellung. In dem Satz *The dog broke the plate* wird *dog* als Agens gesehen, in dem Satz *The hammer broke the plate* ist *hammer* ein Instrument.

Aus den Erfahrungen des trainierten Korpus konnte das Netz bei Sätzen, in denen Subjekt und Objekt semantisch gleichwertig sind, anhand der Satzstellung eine Zuordnung vor-

nehmen. So ist in den Sätzen *The boy bit the girl* und *The girl bit the boy* jeweils das Subjekt der Handelnde.

Ambiguitäten, die ohne Kontextbezug auch für den menschlichen Leser nicht auflösbar sind, waren für das Netz zumindest erkennbar. In *The bat broke the window* wurde *bat* sowohl als Agens als auch als Instrument gesehen, da es in den Trainingszyklen sowohl als Fledermaus als auch als Baseballschläger kodiert war.

Das Netz von McClelland und Kawamoto ist also in der Lage, anhand der wiederholten Betrachtung von Beispielsätzen ein Muster in der Verwendung der einzelnen Worte zu erkennen und bei der Zuordnung der Rollen zu den Konstituenten sowohl die Semantik der Worte als auch die Satzstellung zu berücksichtigen.

III.2 Sequentielle Satzverarbeitung mit KNNs

Ein künstliches neuronales Netz zu sequentiellen Satzverarbeitung wurde 1990 von Elman entwickelt, der durch zwei Zwischenschichten, die in beide Richtungen miteinander verbunden sind, soetwas wie Erinnerungsvermögen simulieren konnte und dadurch sein Netz unter anderem erfolgreich darauf trainierte innerhalb einer Wortkette das nächste Wort vorherzusagen.

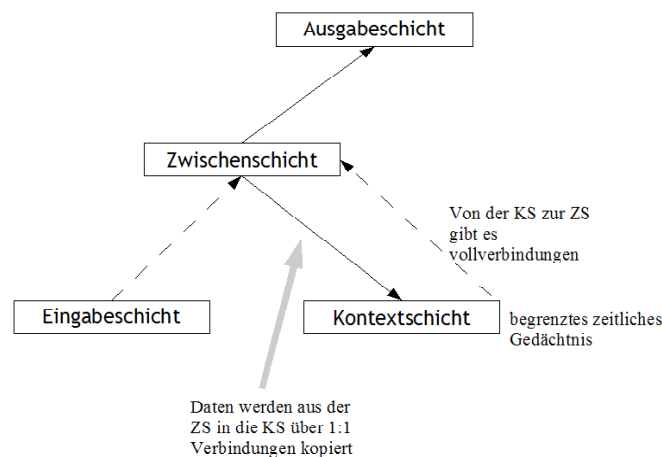


Abbildung III.2.1: Die Kontextschicht ist ein Hopfieldnetz, das vollständig mit der Zwischenschicht verbunden ist

Im Vorfeld wurden aus 29 lexikalischen Einheiten, wie *man*, *woman*, *cat*, *see* 10.000 Zwei- und Drei-Wort-Sätze (*man eat cookie*) generiert, wobei jedes Wort, wie im Netz aus Kapitel III.1, durch einen 31-Bit Merkmalsvektor codiert wurde. Das Netz besitzt eine Ein- und Ausgabeschicht mit 31 Neuronen und die in Abbildung III.2.1 dargestellten zwei inneren Schichten mit jeweils 150 Neuronen. Alle Neuronenaktivierungen, die von der Zwischen-

schicht zur Ausgabeschicht geleitet werden, gehen ebenso zur Kontextschicht, einem Hopfieldnetz, das wiederum vollständig mit der Zwischenschicht verbunden ist.

Über die verallgemeinerte Delta-Regel wurde nun das Netz darauf trainiert, nach der Eingabe eines Wortes das nächste Wort vorherzusagen. Da dies bei Drei-Wort-Sätzen und den zahlreichen Kombinationsmöglichkeiten kaum möglich ist, konnte das Netz nur in den seltensten Fällen die korrekte Codierung für das gesuchte Wort zurück liefern.

Nach dem Trainingszyklus mit den 10.000 Sätzen war aber dennoch eine interessante Beobachtung feststellbar. Die Worte wurden nach 31 Merkmalen kodiert, aber von Elman im Vorfeld in Kategorien, wie Noun-Human, Noun-Anim oder Verb-Intran (für Intransitiv) eingeordnet. Diese Kategorien waren jedoch nicht Bestandteil der Merkmalsvektoren. Nichtsdesto trotz liess sich durch eine Clusteranalyse feststellen, dass Worte die zur gleichen Kategorie gehörten auch die gleichen Neuronen der Kontextschicht verstärkt aktivierten. Das Netz hat also anhand der trainierten Sätze und ihrer Merkmale abstraktere Muster finden können, die der Mensch zum Verständnis der Welt ebenfalls aus der Detailvielfalt heraus erkennt. Insbesondere die Unterscheidung zwischen transitiven und intransitiven Verben zeigt, dass das Netz aufgrund der vielen Beispielsätze für besonders abstrakte Strukturen sensibilisiert werden konnte.

IV Ausblick

In den letzten beiden Kapiteln wurden zum einen Agentensysteme und zum anderen künstliche neuronale Netze vorgestellt, mit denen sich unterschiedliche Aspekte von Sprache simulieren lassen. Während in der Agentensimulation die Entstehung eines kollektiven Wortschatzes anhand von Kommunikationsaufgaben untersucht wurde, zeigen die beiden vorgestellten neuronalen Netze, wie lernende Systeme aussehen können, die Sprechakte durch Erfahrungswerte richtig analysieren und sich dabei interne Strukturen der Netze ergeben, die darauf hinweisen, dass die Mustererkennungsfähigkeiten des Netzes zu abstrakteren Kategorisierungsarten führen.

Im Rahmen meiner Diplomarbeit, die ich im Wintersemester 2006/07 beginnen möchte, plane ich ein Tool zu schreiben, mit dem man Simulationsmodelle erstellen kann, die die Kombination von neuronalen Netzen und Agentensimulationen erlauben, wodurch komplexere Analysen der natürlichen Sprache möglich werden sollen. Die neuronalen Netze sollen dabei als objektorientierte Strukturen im System sehr individuell vom Anwender selbst gebaut werden können. Er kann sich entweder sein eigenes Netz durch Klicken schnell zusammenbauen oder mittels Assistenten grössere und komplexere Netze, wie Feedforward-Netze mit mehreren Schichten oder vollständig vernetzte Hopfieldnetze, von der Anwendung erstellen lassen.

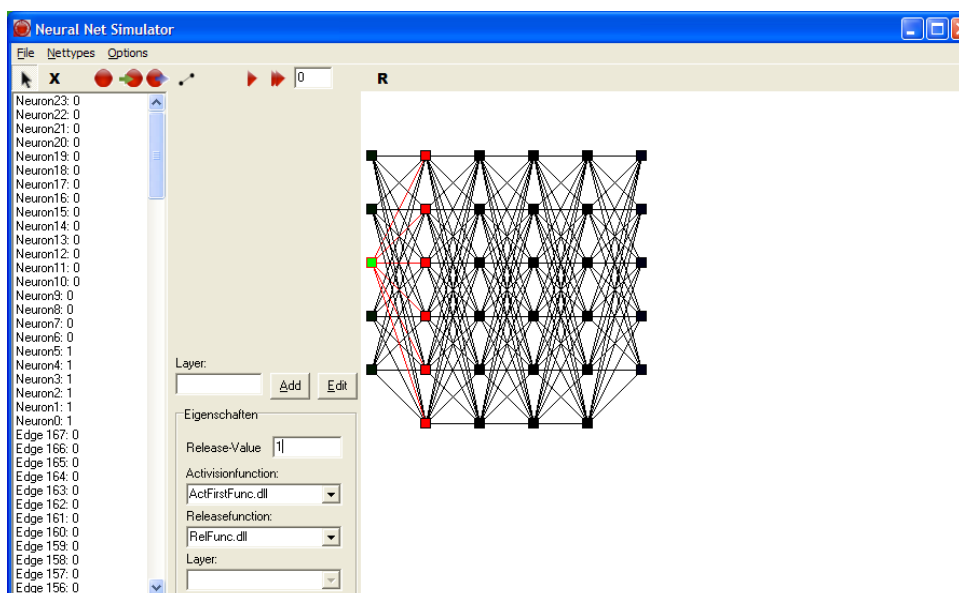


Abbildung IV.1: Im Neural Net Simulator lassen sich Netze schnell zusammenklicken oder komplexe Strukturen über Assistenten erzeugen

Für jedes der Neuronen lassen sich eigene Aktivierungs- und Release-Funktionen wählen, den Kanten lassen sich Gewichte und Übertragungsfunktionen zuweisen. Ein Neuron kann

optional die Rolle eines Sensors oder eines Nervs zugewiesen bekommen. Dieses Konzept soll es ebenfalls in einer zweiten Anwendung geben, in der der Benutzer durch geometrische Formen oder ebenfalls durch Assistenten eigene Agenten zusammenbauen kann. Diesen Agenten können dann die selbst erstellten neuronalen Netze zugewiesen werden. Die kognitiven und motorischen Fähigkeiten des Agenten, also die Sensoren und die Nerven, lassen sich dann mit den entsprechenden Neuronen des neuronalen Netzes verbinden.

Mit Hilfe dieser zweiten Anwendung soll dann der Benutzer schnell und leicht eigene Agentensimulationen entwickeln können, um sprachliche Phänomene untersuchen zu können. Das System soll also das erste seiner Art sein, dass es erlaubt eine grosse Vielzahl von unterschiedlichen Agentensystemen in Kombination mit neuronalen Netzen erstellen zu können. Auf Grund der Tatsache, dass sich durch Ziehen und Klicken Neuronen miteinander von Hand verbinden lassen, können aber auch Zustandsgraphen modelliert werden, die den Agenten zugewiesen werden.

Innerhalb der Planungsphase dieses Projektes ist die wichtigste Herausforderung die nötigen Einschränkungen und Richtlinien zu finden, die ein möglichst breites Spektrum an Agentensimulationen erlauben, aber diese Anwendung nicht zur eierlegenden Wollmilchsau machen, dessen technische Umsetzung dann in einer Sackgasse endet.

Mit der in diesem Projekt entstehenden Software soll vor allen Dingen analysiert werden können, welche Rolle komplexe Handlungs- und Interaktionsmöglichkeiten von Agenten auf die Entwicklung von Sprache haben bzw. wie weit sich die Realität vereinfachen lässt, so dass Kasusgrammatik und kreatives Sprechen immer noch entstehen können.