Wenn Ungenauigkeit von Vorteil ist[1]

Christoph Hesse

Die moderne Welt ist die reinste Zahlenflut: im Beruf, in den Nachrichten, im Sport, bei der Urlaubssuche, bei der Wohnungssuche, selbst in der Freizeit wird gemessen und verglichen, was das Zeug hält. Die meisten Leute beteuern, Präzision sei wichtig, doch betrachtet man den tatsächlichen Gebrauch von Zahlenangaben, finden sich Formulierungen wie *etwa 20*, *mehr als die Hälfte* oder *weniger als 30%* viel häufiger als exakte Angaben. Es gilt die Regel: Nebensächliches wird vernachlässigt. Abstraktion ist allgegenwärtig und sie ist notwendig, um mit der sonst überwältigenden Flut von Zahlen und Informationen zurecht zu kommen.

Schaut man sich an, wie diese ungenauen Zahlenangaben gemeint sind, so stellt man fest, dass zum Beispiel *mehr als 100* eben nicht wie in der Mathematik jede Zahl größer als 100 meint, sondern Sprecher immer davon ausgehen, dass der mögliche Wertebereich nach oben hin begrenzt ist. Die Vermutung liegt nahe, dass solch eine Begrenzung sich darauf zurückführen lässt, dass unsere Welterfahrung uns lehrt, dass Wertebereiche im wahren Leben immer endlich sind und Unendlichkeiten lediglich in der Mathematik existieren.

In der Linguistik sind wir daran interessiert, welche Verallgemeinerungen sich über ein bestimmtes sprachliches Phänomen treffen lassen. Bezogen auf Zahlen, die durch *mehr als* oder *weniger als* modifiziert sind, interessiert uns als Linguisten, ob es Verallgemeinerungen gibt, die über den konkreten Einzelfall hinausgehen und nicht von dem jeweiligen Kontext abhängen, in dem die Modifizierung verwendet wurde. Wir suchen also nach Fallbeispielen, in denen man für die Interpretation von modifizierten Zahlen nicht auf Erfahrungswerte oder Weltwissen zurückgreifen kann, denn anhand dieser Beispiele sollten wir ungetrübt auf das allgemein Gültige blicken.

Um den Ursachen unserer Intuitionen auf die Spur zu kommen, kann es hilfreich sein, Negativbeispiele zu verstehen. Die Äußerung *Berlin hat mehr als 1000 Einwohner* klingt zum Beispiel falsch, weil man ja weiß, dass Berlin viel mehr als nur tausende Einwohner hat. Der Schwellenwert von 1000 ist viel zu niedrig angesetzt. Mit anderen Worten, die Zahl Tausend in Kombination *mit mehr als* impliziert einen Wertebereich, der weit unterhalb der tatsächlichen Einwohnerzahl von Berlin liegt, und diese Annahme bleibt bestehen, selbst wenn man sich unsicher ist, wieviel es tatsächlich sind. Stephanie Solt und Uli Sauerland vom ZAS konnten 2012 zusammen mit Chris Cummins zeigen, dass selbst wenn es Sprechern an Erfahrungswerten fehlt, um den möglichen Wertebereich plausibel einzugrenzen, sie es trotzdem tun. Die Schätzungen verschiedener Leute liegen dabei überraschend nah beieinander, wie ihre Experimente zeigten.

[1] Ein Projekt zusammen mit Anton Benz und gefördert durch die Deutsche Forschungsgemeinschaft (Nr. 232196050).

Woher kommt also die gedankliche Eingrenzung des Wertebereichs, wenn nicht durch Erfahrungswerte? Die Lösung, die Stephanie Solt, Chris Cummins und Uli Sauerland damals vorschlugen, war folgende: Hat man keine Erfahrungswerte, auf die man zurückgreifen kann, muss man Annahmen über den Wertebereich an der genannten Zahl selbst fest machen. Was bei einer systematischen Betrachtung schnell ins Auge fällt ist, dass Beispiele mit runden Zahlen wesentlich häufiger auftreten als Beispiele mit exakten Zahlen: *mehr als 100* kommt eher vor als *mehr als 107,4*. Warum wurde gerundet? Weil Sprecher mit der Formulierung *mehr als 100* für 114,3 suggerieren, dass der Unterschied zwischen 100 und 114,3 in diesem konkreten Fall vernachlässigt werden kann. Umgekehrt lässt die Rundheit der Zahl aber auch Rückschlüsse darauf zu, wie stark gerundet wurde und kann damit indirekt Aufschluss über den vermuteten Wertebereich geben. Ausnahmen bilden Fälle, in denen bestimmte Grenzwerte eine Rolle spielen. Zum Beispiel, wenn bei einer Abstimmung eine bestimmte Zahl von Stimmen für eine Mehrheit notwendig ist, findet man Formulierungen wie *Es sind mehr als 327 Stimmen notwendig* obwohl 327 keine runde Zahl ist.

Nehmen wir an, wir wissen, der tatsächliche Wert ist 114,3. Wenn wir runden möchten, müssen wir uns zuerst entscheiden, wie stark wir runden wollen. Auf Einer gerundet wird aus 114,3 die Zahl 114, denn 0,3 ist weniger als 0,5 — liegt also in der unteren Hälfte des Zahlenbereiches von Null bis Eins — und wir runden deshalb ab statt auf. Man würde also *mehr als 114* verstehen als „zwischen 114 und 115.“ Wollen wir 114,3 auf Zehner runden, so runden wir ab, weil 4 unter 5 liegt — also in der unteren Hälfte des Zahlenbereiches Null bis Zehn. Man würde daraus schließen, dass der tatsächliche Wert von *mehr als 110* zwischen 110 und 120 liegt, was auf 114,3 zutrifft. Runden wir auf Hunderter, so könnten wir *mehr als 100* sagen, denn 114,3 ist näher an 100 als 200. Je stärker wir runden, desto weniger präzise sind unsere Angaben. Im Umkehrschluss bedeutet das, dass wir annehmen können, dass eine stärkere Rundung zulässig ist, wenn Präzision eine untergeordnete Rolle spielt.

Stephanie Solt, Chris Cummins und Uli Sauerland vermuteten, dass bei runderen Zahlen der Wertebereich, welcher den erwarteten tatsächlichen Wert enthält, mit der Rundheit wächst. Nehmen wir als Beispiel an, wir fragen nach der nächsten Tankstelle und bekommen als Antwort eine Entfernungsangabe mit entweder *mehr als 26 km* oder *mehr als 25 km*. Bei *mehr als 26 km* muss die tatsächliche Entfernung über 26 km liegen, aber wo genau, wissen wir nicht. Da 26 jedoch weniger rund ist als 25, liegt die Vermutung nahe, dass bei 26 km auf volle Kilometer gerundet wurde, die tatsächliche Entfernung also zwischen 26 und 27 km liegt. *Mehr als 25 km* hingegen suggeriert, dass auf das nächstniedrigere 5 Kilometer-Intervall gerundet wurde, der tatsächliche Wert also zwischen 25 und 30 km liegt. Wenn man nun noch präzisere Vermutungen anstellen will, kann man sich vor Augen führen, dass, wenn die tatsächliche Entfernung näher an 27 km als an 26 km liegen würde, der Sprecher oder die Sprecherin sicherlich eher *weniger als 27 km* gesagt hätte. *Mehr als 26 km* suggeriert also nicht nur, dass die Entfernung zwischen 26 und 27 km ist, sondern auch, dass sie näher an 26 als an 27 km liegt, also zwischen 26 und 26,5 km. Die gleichen Überlegungen könnte man für *mehr als 25 km* anstellen und käme zu dem Schluss, dass die tatsächliche Entfernung näher an 25 km als an 30 km sein muss, also zwischen 25 und 27,5 km. Wir sehen, nicht nur ist der

Wertebereich, den wir für *mehr als 25 km* und *mehr als 26 km* in Betracht ziehen, unterschiedlich groß, sondern der vermutete tatsächliche Wert rückt bei *mehr als 25 km* weiter von 25 km weg als bei *mehr als 26 km* von 26 km.

Bei der Entwicklung eines Dialogsystems, welches Zahlenangaben mit *mehr als* und *weniger als* generiert, stellten wir, Anton Benz und Christoph Hesse, fest, dass die Rundheitsregel von Cummins, Sauerland & Solt bei größeren Zahlen nicht die vorhergesagten Erwartungen bei Anwendern weckte. Deshalb machten wir uns auf die Suche nach einer alternativen Erklärung. Wir begannen systematisch Zahlen in Zehnerschritten in jener Größenordnung experimentell zu testen, die auch Cummins, Sauerland & Solt untersucht hatten. Dazu zeigten wir Probanden kurze Szenarien und ließen sie Schätzwerte abgeben, wie im folgenden Beispiel:

> You overhear the following conversation between Brooke and Oliver.
>
> Brooke: I heard there’s even a petition against the highway construction project.
>
> Oliver: [modified numeral, e.g. “more than 100”] people already signed it.
>
> How many people do you think signed the petition?
> Between _____ and _____, most likely _____.

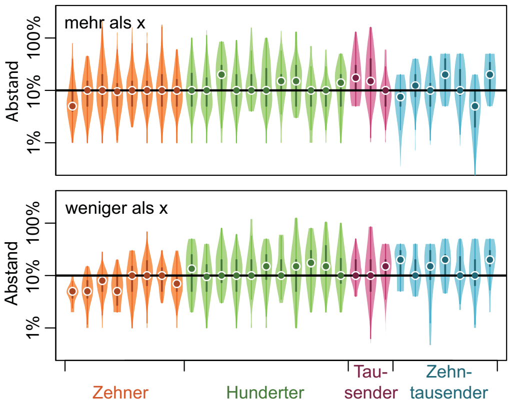
Sowohl in den Experimenten von Cummins, Sauerland & Solt als auch in unseren Experimenten zeigte sich, dass Probanden nicht alle Werte, die sie im Wertebereich in Betracht ziehen, für gleich wahrscheinlich halten. Viel mehr haben sie eine vage Vorstellung davon, in welcher Größenordnung sich der tatsächliche Wert von *mehr als x* oder *weniger als x* bewegen könnte, und stufen die Wahrscheinlichkeit der möglichen Werte im Verhältnis zu ihrem erwarteten tatsächlichen Wert ab. Mit den ersten zwei Schätzwerten in obigen Beispiel geben Probanden an, wie groß der Wertebereich ihrer Meinung nach ist. Mit dem dritten Schätzwert geben sie an, welchen Wert innerhalb des möglichen Wertebereichs sie für am wahrscheinlichsten halten. Es stellte sich zu unserer Überraschung heraus, dass der geschätzte wahrscheinlichste Wert und die zugrunde gelegte Länge des Intervalls nicht durch die Rundheit von x bestimmt wird.

Die Tabelle in A zeigt, welchen möglichen Wertebereich Experimentteilnehmer für verschiedene Zahlen x modifiziert durch *mehr als x* und *weniger als x* jeweils in Betracht ziehen. Der je x aufgelistete Grenzwert ist der Grenzwert, der am häufigsten in Experimenten genannt wurde. Es fällt auf, dass der von x weggewandte Grenzwert für benachbarte x oftmals identisch ist. Zum Beispiel wird für *weniger als 60*, *weniger als 70* und *weniger als 80* meist der Grenzwert 50 angegeben; für *weniger als 90* und *weniger als 100* ist der häufigste Grenzwert 80; für *weniger als 110* bis *weniger als 170* ist er 100, für *weniger als 180* bis *weniger als 200* dann 150. Es ist weiterhin auffällig, dass die genannten Grenzwerte selbst runde Zahlen sind, welche das betreffende Intervall logisch unterteilen, so zum Beispiel das Intervall zwischen 0 und 100 mit dem Grenzwert 50. Gleichermaßen halbiert 150 das Intervall zwischen 100 und 200, 500 das Intervall zwischen 0 und 1000, 15.000 das Intervall zwischen 10.000 und 20.000. Die Größe des in Betracht gezogenen Wertebereichs ist also mal zunehmend, mal abnehmend, abhängig davon, wie nah x auf dem Zahlenstrahl zu seinem nächsten möglichen Grenzwert steht. Aus diesem Grund konnten wir mit unseren Experimenten auch nicht die von Cummins, Sauerland & Solt vermutete Korrelation zwischen Rundheit von x und Größe des möglichen Wertebereichs bestätigen. In unseren Experimenten zeigte sich, nicht die Rundheit von x ist entscheidend, sondern der Abstand zur nächstrunderen Zahl y.

A Erwarteter Wertebereich

weniger als x		mehr als x	
x	Grenzwert	x	Grenzwert
20	10	20	30
30	20	30	50
40	30	40	60
50	40	50	100
60	50	60	100
70	50	70	100
80	50	80	100
90	80	90	100
100	80	100	120
110	100	110	120
120	100	120	150
130	100	130	150
140	100	140	200
150	100	150	200
160	100	160	200
170	100	170	200
180	150	180	200
190	150	190	200
200	150	200	250
1000	500	1000	1500
1100	1000	1100	1500
1400	1000	1400	1500
15.000	10.000	15.000	20.000
16.000	10.000	16.000	20.000
19.000	15.000	19.000	20.000
20.000	15.000	20.000	25.000
21.000	15.000	21.000	25.000
24.000	20.000	24.000	25.000
25.000	20.000	25.000	30.000

B Abstand des für x erwarteten Werts relativ zur Größenordnung von x



Bezüglich des erwarteten tatsächlichen Werts von *mehr als x* und *weniger als x* würde man erwarten, dass er weiter von x wegrückt je runder x ist. In den beiden Diagrammen in B sehen wir jedoch, dass unabhängig davon, ob x von der Größenordnung her im Zehnerbereich (x = 20, 30, 40, … 90) oder im Hunderterbereich (100, 200, …) oder im Tausenderbereich (1000, 1100, …) oder im Zehntausenderbereich (10.000, 15.000, 20.000, 25.000, …) liegt, die meisten Probanden annahmen, dass der tatsächliche Wert ca. 10% von x entfernt ist, relativ zur Größenordnung von x. Wenn x also im 3-stelligen Bereich ist, erwarten sie, dass der tatsächliche Wert etwa 10% von 100 = 10 entfernt von x liegt; im 4-stelligen Bereich erwarten sie einen Abstand von 100; im 5-stelligen Bereich einen Abstand von 1000. Nur im 2-stelligen Bereich wichen sie davon ab und nahmen auch wie im 3-stelligen Bereich einen Abstand von 10 an. Entgegen der Diagnose von Cummins, Sauerland & Solt beruhen Vermutungen über den tatsächlichen Wert also weniger auf der Rundheit von x, sondern viel mehr auf seiner Größenordnung. Dieses Mitskalieren mit der Größenordnung ist ein kognitives Phänomen, welches schon seit über hundert Jahren in der Psychologie als Weber-Fechner-Regel bekannt ist.

An diesem Punkt wurde uns klar, dass wir in unseren Experimenten auf eine Verknüpfung von Erkenntnissen aus der Linguistik mit Erkenntnissen aus Psychologie und Kognitionswissenschaften gestoßen waren: dem sogenannten *Approximate Number Sense*. Der Approximate Number Sense ist ein beim Menschen (und anderen Tieren) angeborener, unbewusster kognitiver Mechanismus, mit dem Mengenunterschiede subjektiv wahrgenommen werden. Charakteristisch für den Approximate Number Sense ist, dass die Auflösung, mit der Unterschiede wahrgenommen werden können, mit der Größenordnung der Menge abnimmt — dieselbe Weber-Fechner-Regel, die wir bezüglich des erwarteten tatsächlichen Wertes in unseren Experimenten sahen. Bei einem Workshop, welchen wir zum Thema „The Meaning of Numerals: cognitive, experimental, and semantic perspectives" am ZAS ausrichteten, konnten wir Vertreter aus diesen unterschiedlichen Disziplinen zusammenbringen.

Diese fachbereichsübergreifende Perspektive erlaubte uns, unsere Ergebnisse auf eine verblüffend einfache Formel zu bringen: Probanden erwarten, dass der tatsächliche Wert innerhalb der Auflösungsungenauigkeit liegt, mit der Größenunterschiede mit dem Approximate Number Sense wahrgenommen werden können. Der mögliche Wertebereich bewegt sich maximal in der gleichen Größenordnung der genannten Zahl. Da die Genauigkeit des Approximate Number Sense mit zunehmender Größenordnung der Zahl abnimmt, wird proportional auch der Abstand zwischen der genannten Zahl und dem vermuteten tatsächlichen Wert immer größer.

Am Ende halfen uns die Erkenntnisse aus dieser Studie dabei, Zahlenangaben in unserem Dialog-System zu vereinfachen, ohne dabei Anwender zu belügen, zu überrumpeln oder mit der Zahlenflut zu überfordern. Es zeigt sich, das richtige Maß an „Ungenauigkeit" kann von Vorteil sein.

Cummins, C., Sauerland, U., & Solt, S. (2012). Granularity and scalar implicature in numerical expressions. *Linguistics and Philosophy*, 35, 135–169. https://doi.org/10.1007/s10988-012-9114-0.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Revised and Updated ed.). USA: Oxford University Press.

Fechner, G. T. (1860). *Elemente der Psychophysik*, vol. 2. Leipzig: Breitkopf und Härtel.