

*Johannes Akkach (Heidelberg), Dieter Friedrich (Trier), Meike Miller (Heidelberg),  
Vanessa Schulz (Frankfurt am Main)*

**Texte der Doppelbesteuerungsabkommen  
in der CAT-gestützten Übersetzung**



**Editors:**

Viktorija Bilić

Anja Holderbaum

Anne Kimmes

Joachim Kornelius

John Stewart

Christoph Stoll

**Publisher:**

Wissenschaftlicher Verlag Trier

Johannes Akkach (Heidelberg), Dieter Friedrich (Trier), Meike Miller (Heidelberg),  
Vanessa Schulz (Frankfurt am Main)

## Texte der Doppelbesteuerungsabkommen in der CAT-gestützten Übersetzung

### Abstract:

Ziel der Untersuchung ist es, festzustellen, inwieweit sich herkömmliche Programme für die computergestützte Übersetzung (CAT) bei der Fachübersetzung von steuerrechtlichen Texten im Sinne einer Prozessoptimierung und Qualitätssicherung nutzen lassen. Dazu wird ein CAT-System per Alignment vorbereitet: Ein bereits existierender Steuerfachtext wird in zwei Sprachen eingelesen und seine Übersetzungseinheiten miteinander verknüpft, damit diese für ähnliche Passagen im zu übersetzenden Text als Vorlage dienen können. Als möglicher Text für das Alignment werden zum einen das allgemeine OECD-Musterabkommen und zum anderen das länderspezifische deutsch-kanadische Doppelbesteuerungsabkommen (DBA) betrachtet, jeweils auf Deutsch und auf Englisch. Für die praktische Durchführung wird das CAT-System *Across* gewählt. Es gilt festzustellen, wie die standardmäßige satzbasierte Segmentierung des Programms für eine hypotaktische Textstruktur und eine Struktur der Form Artikel-Paragrafen-Subparagrafen funktionalisiert werden kann. Weiter wird untersucht, ob bei der Vorbereitung vor dem Alignment eine Anpassung des zugrunde liegenden Textes sinnvoll ist.

This investigation seeks to establish how effective computer-assisted translation programs (CAT) prove to be when used in the technical translation of tax-related documents, and whether such programs enhance their professional quality. The analyzed documents originated from the OECD-Master Agreement and deal with German-Canadian Double-Taxation (DBA), and thus they appear accordingly in German and in English. The *Across* CAT-system was selected for its practicality, and was investigated to determine whether it is a feasible option for use in tax-related translations involving complex sentence structures in paragraph form.

**Keywords:** Translation-Memory-Systeme; CAT; Steuerabkommen; Fachsprache; Segmentierung

### Inhalt:

1	Doppelbesteuerungsabkommen als Fachtexte des internationalen Steuerrechts: Eine Orientierung .....	2
2	Zielsetzungen: Doppelbesteuerungsabkommen und die Übersetzung mit CAT-Systemen .....	3
3	Die Auswahl eines geeigneten Abkommens für die Erstellung des Translation Memorys.....	4
4	Das Anpassen der Texte vor dem Einlesen und das Einlesen des gewählten Abkommens .....	11

5	Die Konkordanzsuche im Translation Memory und die Einbindung einer Terminologiedatenbank .....	17
5.1	Die Konkordanzsuche .....	17
5.2	Die Terminologiedatenbank .....	18
6	Fazit .....	21
7	Bibliographie .....	22

## **1 Doppelbesteuerungsabkommen als Fachtexte des internationalen Steuerrechts: Eine Orientierung**

Der Aufbau und die Pflege weltweiter Wirtschaftsbeziehungen und die Gründung von Auslandsniederlassungen ist heute nicht mehr allein den großen Wirtschaftskonzernen vorbehalten, vielmehr betätigen sich auch zunehmend kleinere und mittelständische Unternehmen auf den internationalen Märkten. Aus diesem Status eines „Global Players“ resultieren beachtliche wirtschaftliche Vorteile. Gleichzeitig werden diese Unternehmen aufgrund ihrer grenzüberschreitenden Tätigkeit mit dem Steuerrecht anderer Nationen konfrontiert und können unter Umständen unter die Steuerhoheit dieser Staaten fallen. Die unterschiedliche Beurteilung der wirtschaftlichen Aktivitäten im Kontext bestehender verschiedenartiger Steuersysteme kann zum einen dazu führen, dass Geschäftsvorfälle sowohl im Inland als auch im Ausland der Besteuerung und damit einer Mehrfachbesteuerung unterliegen. Zum anderen kann jedoch auch durch bestehende Gestaltungsspielräume eine niedrigere Steuerbelastung oder gar Nichtbesteuerung erreicht werden.

Aufgabe des internationalen Steuerrechts ist es, diese Doppel- oder Mehrfachbesteuerung bzw. eine doppelte Nichtbesteuerung zu vermeiden sowie die Steuerfolgen der unterschiedlichen Steuerrechtsordnungen an international gültigen Besteuerungsprinzipien auszurichten und aufeinander abzustimmen (vgl. Tipke und Lang 2013).

Das internationale Steuerrecht ist kein einheitlich normiertes Gesetz, sondern umfasst die Gesamtheit aller Normen, die die Besteuerung grenzüberschreitender Aktivitäten regeln (vgl. Brähler 2012). Es basiert auf dem Völkerrecht, hier vor allem auf den völkerrechtlichen

Verträgen, auf supranationalem Recht, insbesondere dem Europarecht und einer Vielzahl von Einzelnormen in unterschiedlichen nationalen Steuergesetzen.

Die nationalen Normen im Zusammenhang mit der Besteuerung auslandsbezogener Sachverhalte werden als Außensteuerrecht bezeichnet (vgl. [Tipke und Lang 2013](#)). Sie beinhalten i. d. R. kollisionsbegründende Maßnahmen, das heißt, sie begründen nationale Steueransprüche, die sich mit Normen ausländischer Rechtsordnungen überschneiden und damit eine Doppelbesteuerung auslösen (vgl. [Brähler 2012](#)). Die Vermeidung der Doppelbesteuerung erfolgt in erster Linie durch den Abschluss von Doppelbesteuerungsabkommen (DBA). Hierbei handelt es sich um völkerrechtliche, bilaterale Verträge.

## **2 Zielsetzungen: Doppelbesteuerungsabkommen und die Übersetzung mit CAT-Systemen**

Im Folgenden wird untersucht, ob und inwieweit sich marktübliche Programme für die computergestützte Übersetzung (CAT) bei der Übersetzung von steuerrechtlichen Texten einsetzen lassen und geprüft, ob sich damit eine Prozessoptimierung und eine Qualitätssicherung verbindet.

Im Kontext der Übersetzung handelt es sich bei den DBA um eine spezielle Textsorte, die jeweils in beiden Sprachen in beiden Ländern als Gesetzestext verfügbar ist. Als Grundlage eines solchen Abkommens kann das von der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) erarbeitete Musterabkommen zur Doppelbesteuerung dienen, das in englischer Sprache verfügbar ist ([OECD 2010](#)).

Diese Texte charakterisiert eine ausgeprägte hypotaktische Satzstruktur und eine Textstruktur der Form Artikel-Paragraph-Subparagraph. Es wird zu prüfen sein, wie ein satzbasiertes CAT-System für diese Eigenschaften des Fachtextes funktionalisiert werden kann.

In einem angenommenen Arbeitsszenario gilt es, ein CAT-System durch Einlesen bestehender Doppelbesteuerungsabkommen in Ausgangs- und Zielsprache vorzubereiten. Hierbei wird mit dem CAT-System *Across* in der Version 5.0 gearbeitet.

### **3 Die Auswahl eines geeigneten Abkommens für die Erstellung des Translation Memorys**

Als erster methodischer Schritt des zuvor genannten Arbeitsszenarios werden die beiden bestehenden Fassungen eines DBA per Alignment in das Translation Memory (TM) importiert. Damit ist es möglich, während der späteren Übersetzung in dieser Sprachkombination auf die vorher gespeicherten Übersetzungen des Alignments zurückzugreifen. Diese liegen in der Praxis im TM-System (hier synonym zu „CAT-System“ verwendet) in Form von übersetzten Satzpaaren vor und können für neue Übersetzungen als fachsprachliche Vorlagen im Ausgangstext und Zieltext verwendet werden. Es werden zwei korrespondierende Fassungen eines DBA in das TM von *Across* eingepflegt. Hierbei soll zunächst untersucht werden, ob für den Übersetzungsprozess das OECD-Musterabkommen oder eine länderspezifische Version zwischen Deutschland und einem bestimmten englischsprachigen Vertragsstaat geeigneter ist.

Als Grundlage für ein länderspezifisches Vertragsverhältnis wird von offizieller Seite in der Regel das OECD-Musterabkommen herangezogen. Bei der Anpassung des englischsprachigen OECD-Musterabkommens werden üblicherweise neue Eintragungen nicht nur an den vorgesehenen Leerstellen eingefügt, sondern innerhalb der Musterartikel auch die für das jeweilige bilaterale Vertragsverhältnis relevanten Zusatzbestimmungen ergänzt. Im Anschluss ist entweder das vollständige Abkommen mit seinen spezifischen Anpassungen in die relevanten Sprachen der Vertragsstaaten zu übersetzen, oder es sind - im Fall von späteren Änderungen - allein die neu eingefügten Passagen zu übersetzen. Ein länderspezifisches Abkommen auf der Grundlage des OECD-Musterabkommens enthält somit deutlich mehr Text und detailliertere Formulierungen als das Musterabkommen selbst. In Abbildung 1 wird Artikel 2 des OECD-Musterabkommens gezeigt. Dort sind an den relevanten Stellen zum einen generische Bezeichnungen wie „Contracting State“ aufgeführt,

zum anderen werden für die spezifischen Steuern nur Auslassungspunkte als Platzhalter gesetzt.

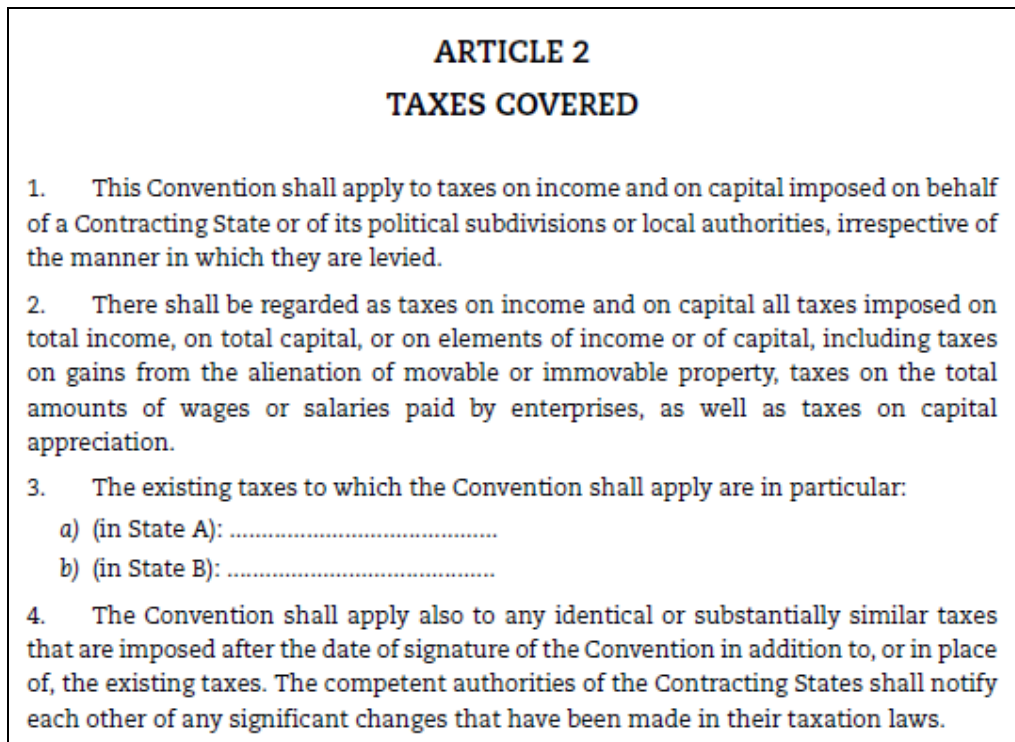


Abbildung 1: Artikel 2 des OECD-Musterabkommens

In den länderspezifischen Abkommen sind hingegen die Anmerkungen zu den deutschen Bundesländern als Einschub aufgeführt, wie in Abbildung 2 bis Abbildung 4 zu sehen ist. Diese in den nachfolgenden Abbildungen umrahmten länderspezifischen Ergänzungen sind im OECD-Musterabkommen nicht enthalten.

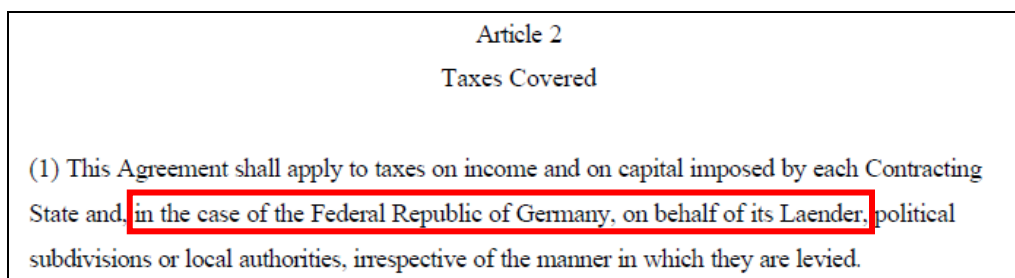


Abbildung 2: Artikel 2, Paragraph 1 des DBA zwischen der Bundesrepublik Deutschland und Kanada

Article 2  
Taxes covered

(1) This Convention shall apply to taxes on income and on capital imposed on behalf of a Contracting State, of a "Land" or a political subdivision or local authority of a "Land" or a Contracting State, irrespective of the manner in which they are levied.

Abbildung 3: Artikel 2, Paragraph 1 des DBA zwischen der Bundesrepublik Deutschland und dem Vereinigten Königreich Großbritannien und Nordirland

Article 2  
Taxes Covered

(1) This Agreement shall apply to taxes on income and on capital imposed on behalf of a Contracting State, of a Land or a political subdivision or a local authority thereof, irrespective of the manner in which they are levied.

Abbildung 4: Artikel 2, Paragraph 1 des DBA zwischen der Bundesrepublik Deutschland und der Republik Singapur

Zusätzlich liefert die englischsprachige Aufzählung durch die Verdeutlichung der deutschen Steuer in ihrer zielsprachlichen Benennung wertvolle fachsprachliche Terminologie von offizieller Seite. Das OECD-Musterabkommen hingegen enthält an den zur Bearbeitung vorgesehenen Stellen allein inhaltliche Platzhalter, die in einem Translation Memory keinen Mehrwert ergeben.

3. The existing taxes to which the Convention shall apply are in particular:  
a) (in State A): .....  
b) (in State B): .....

Abbildung 5: Artikel 2, Paragraph 3 des OECD-Musterabkommens

Der Unterschied ist im Folgenden in Abbildung 6 bis Abbildung 8 zu sehen, in denen Paragraph 3 der länderspezifischen Abkommen mit Kanada, dem Vereinigten Königreich Großbritannien und Nordirland sowie der Republik Singapur gezeigt wird:

- (3) The existing taxes to which the Agreement shall apply are:
- a) in the case of Canada:
- the taxes imposed by the Government of Canada under the *Income Tax Act*, (hereinafter referred to as "Canadian tax");
- b) in the case of the Federal Republic of Germany:
- aa) the income tax (Einkommensteuer),  
 bb) the corporation tax (Körperschaftsteuer),  
 cc) the capital tax (Vermögensteuer),  
 dd) the trade tax (Gewerbsteuer), and  
 ee) the solidarity surcharge (Solidaritätszuschlag),  
 (hereinafter referred to as "German tax").

Abbildung 6: Artikel 2, Paragraph 3 des DBA zwischen der Bundesrepublik Deutschland und Kanada

- (3) The existing taxes which are the subject of this Convention are in particular:
- a) in the Federal Republic of Germany:
- aa) the income tax ("Einkommensteuer");  
 bb) the corporation tax ("Körperschaftsteuer");  
 cc) the trade tax ("Gewerbsteuer"); and  
 dd) the capital tax ("Vermögensteuer");  
 including the supplements levied thereon  
 (hereinafter referred to as "German tax");
- b) in the United Kingdom:
- aa) the income tax;  
 bb) the corporation tax; and  
 cc) the capital gains tax;  
 (hereinafter referred to as "United Kingdom tax").

Abbildung 7: Artikel 2, Paragraph 3 des DBA zwischen der Bundesrepublik Deutschland und dem Vereinigten Königreich Großbritannien und Nordirland



(3) The existing taxes to which this Agreement shall apply are in particular:

a) in the Federal Republic of Germany:

the income tax (Einkommensteuer),  
the corporation tax (Körperschaftsteuer),  
the capital tax (Vermögensteuer), and  
the trade tax (Gewerbesteuer)  
including the supplements levied thereon  
(hereinafter referred to as "German tax");

b) in Singapore:

the income tax  
(hereinafter referred to as "Singapore tax").

Abbildung 8: Artikel 2, Paragraph 3 des DBA zwischen der Bundesrepublik Deutschland und der Republik Singapur

Die englischen Benennungen der deutschen Steuerarten sind sich jeweils ähnlich. Daher kann ein im Translation Memory vorhandenes Segment aus einem der Abkommen bei der Übersetzung der anderen Abkommen erneut genutzt werden. Aus diesem Grund wird im Folgenden ein länderspezifisches DBA und nicht das OECD-Musterabkommen für die Erstellung des Translation Memory per Alignment genutzt, beispielhaft das DBA zwischen Deutschland und Kanada.

Das Einpflegen der Übersetzungseinheiten geschieht durch ein Alignment des deutschsprachigen und des englischsprachigen DBA, in diesem Fall der DBA von Deutschland und Kanada. Dabei wird der Gesamttext automatisch in korrespondierende Absätze und weiter in Satzstrukturen gegliedert (Segmentierung) und diese Segmente einander gegenübergestellt. Dem Übersetzer obliegt die Zuordnung der ausgangs- und zielsprachlich übereinstimmenden Segmente (Sätze). Damit entsteht ein methodisches Problem, das darin gesehen werden kann, dass die Einheiten der formalen Gliederungsstrukturen von DBA andere sind als die anderer Fachtexte. Die hypotaktische Textur der DBA wird durch eine

spezifische Struktur der Form Artikel-Paragrafen-Subparagrafen bestimmt. Dies lässt sich am nachstehenden Textbeispiel des ausgewählten DBA in Abbildung 9 zeigen:

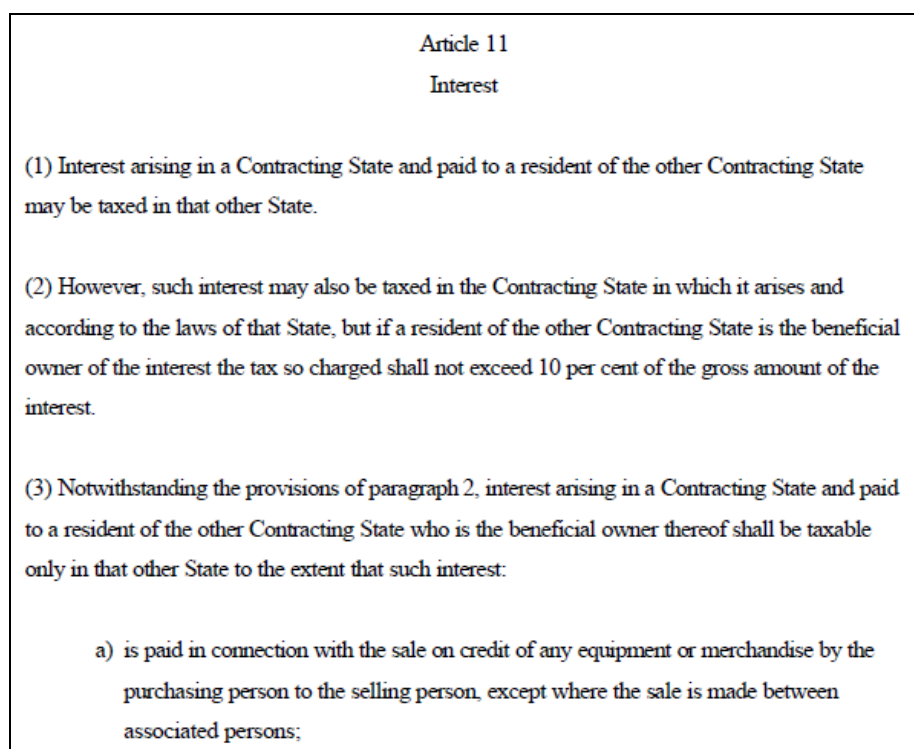


Abbildung 9: Artikel 11 des DBA zwischen der Bundesrepublik Deutschland und Kanada  
(englische Fassung, Auszug)

Im CAT-System wird üblicherweise aufgrund einer zuvor festgelegten Prozentzahl nach Übereinstimmungen des zu übersetzenden Segments mit einem gespeicherten, bereits übersetzten Segment gesucht. Diese beträgt beim CAT-System *Across* standardmäßig mindestens 50 Prozent. Diese Prozentzahl lässt sich zwar individuell nach Belieben anpassen, aber nicht unter 50 Prozent absenken. In den steuerrechtlichen Texten erstrecken sich die hypotaktischen, überlangen Satzstrukturen über mehrere Absätze hinweg. Dies kann dazu führen, dass ein Absatz nicht mehr aus verschiedenen Segmenten, sondern nur noch aus einem großen Segment ohne weitere Strukturierung besteht. Im besten Fall entspricht ein Satz einem Absatz wie in der oben stehenden Abbildung unter Paragraph 2. In diesem Fall werden im Translation Memory vollständige Satzpaare hinterlegt. Dies zeigt die nachstehende Abbildung, hier am Beispiel des deutschen DBA:

Art. 2 Unter das Abkommen fallende Steuern (Art. 2 MA)	
(1) Dieses Abkommen gilt, ohne Rücksicht auf die Art der Erhebung, für Steuern vom Einkommen und vom Vermögen, die für Rechnung eines der beiden Vertragsstaaten, und im Falle der Bundesrepublik Deutschland, eines seiner Länder oder einer ihrer Gebietskörperschaften erhoben werden.	
(2) Als Steuern vom Einkommen und vom Vermögen gelten alle Steuern, die vom Gesamteinkommen, vom Gesamtvermögen oder von Teilen des Einkommens oder des Vermögens erhoben werden, einschließlich der Steuern vom Gewinn aus der Veräußerung beweglichen oder unbeweglichen Vermögens sowie der Steuern vom Vermögenszuwachs.	

Abbildung 10: Übereinstimmung von Satzgrenze und Absatzgrenze

In anderen Fällen setzt jedoch das TM-System beim Alignment die Segmentgrenzen innerhalb von hypotaktischen Sätzen. Diese Fälle ergeben sich, wenn die Autoren des Ausgangstextes durch Zeilenwechsel hypotaktische Sätze strukturieren. Solche Zeilenwechsel innerhalb eines Satzes werden vom TM-System als Satzende und somit als Segmentgrenze interpretiert. Dies hat zur Folge, dass bei der Suche nach vorhandenen Übersetzungen untersatzmäßige Einheiten wie Phrasen und Teilsätze sowie nichtsprachliche Strukturierungselemente wie Nummerierungsziffern oder Einzelbuchstaben als Segmente präsentiert werden. Eine solche Fragmentierung sieht man im folgenden Beispiel aus dem DBA zwischen Deutschland und Kanada:

(3) Zu den bestehenden Steuern, für die das Abkommen gilt, gehören:	
a)	
in Kanada:	
die von der kanadischen Regierung nach dem Einkommensteuergesetz (Income Tax Act) erhobenen Steuern (im Folgenden als „kanadische Steuer“ bezeichnet);	
b)	
in der Bundesrepublik Deutschland	
aa)	
die Einkommensteuer,	

Abbildung 11: Untersatzmäßige Einheiten in der Segmentierung

Dabei kann auch der Fall eintreten, dass die Bestandteile einer normgetragenen Wortkombination, einer Kollokation, in der Ausgangs- oder der Zielsprache auf verschiedene alignierte Segmente verteilt werden. Im nachstehenden Beispiel steht die Basis der Kollokation „Gewinne beziehen“ (Gewinne) im ersten, einleitenden Teil des Absatzes. Der Kollokator (beziehen) steht jedoch davon getrennt nach dem Einschub zweier Unterabschnitte a und b am Beginn eines neuen Segments. Dabei werden die Basis und der Kollokator in voneinander getrennten Datensätzen im Translation Memory gespeichert. Dies

führt dazu, dass die Kollokation als fachsprachliche Einheit und als Stilmittel des Fachtextes neutralisiert wird.

<p>(4) <b>Gains derived</b>, by a resident of a Contracting State from the alienation of:</p> <p>a) shares (other than shares listed on an approved stock exchange in the other Contracting State) forming part of a substantial interest in the capital stock of a company which is a resident of that other State the value of which shares is derived principally from immovable property situated in that other State; or</p> <p>b) an interest in a partnership, trust or estate the value of which is derived principally from immovable property situated in the other Contracting State,</p> <p>may be taxed in that other State. For the purposes of this paragraph, the term "immovable property" does not include property (other than rental property) in which the business of the</p>	<p>(4) <b>Gewinne</b>, die eine in einem Vertragsstaat ansässige Person aus der Veräußerung von</p> <p>a) Anteilen (mit Ausnahme von Anteilen, die an einer anerkannten Börse des anderen Vertragsstaats notiert sind), die zu einer wesentlichen Beteiligung am Kapital einer im anderen Staat ansässigen Gesellschaft gehören und deren Wert überwiegend auf in diesem anderen Staat gelegenen unbeweglichen Vermögen beruht, oder</p> <p>b) Beteiligungen an einer Personengesellschaft, Rechte an einem Trust oder an einem Nachlass, deren Wert überwiegend auf in dem anderen Vertragsstaat gelegenen unbeweglichen Vermögen beruht,</p> <p><b>bezieht</b> können in diesem anderen Staat besteuert werden. Für die Zwecke dieses Absatzes umfasst der Ausdruck "unbewegliches Vermögen" keine Grundstücke (mit Ausnahme von</p>
---	--

Abbildung 12: Darstellung einer Kollokation innerhalb eines Segmentes im englischen Translation Memory und in zwei unterschiedlichen Segmenten im deutschen Translation Memory

Dadurch wird das TM-System bei einer Situation wie in Abbildung 12 dargestellt zwei Segmente anzeigen, die inhaltlich und formal nicht zueinander passen. Im Englischen ist die vollständige Kollokation zu sehen, im Deutschen fehlt der Kollokator.

#### 4 Das Anpassen der Texte vor dem Einlesen und das Einlesen des gewählten Abkommens

Mit Blick auf eine Problemlösung bietet sich die Möglichkeit an, den vorliegenden Text vor dem Einpflegen in das TM-System auf die spezifischen Text- und Layoutstrukturen hin voreinzurichten, um die eingangs erwähnte Vorlage für spätere Übersetzungen zu schaffen.

Um das in Abbildung 12 gezeigte Problem zu umgehen, können vor dem Einlesen und dem Alignment die entsprechenden Zeilenwechsel entfernt werden, damit keine Segmentgrenzen innerhalb von Sätzen gezogen werden und somit keine untersatzmäßigen Einheiten in das Translation Memory gelangen.

Da die einzulesenden Texte zumeist umfangreich sind, sollte aus Gründen der Zeitersparnis für diese Operation eine automatische Erkennung unerwünschter Zeilenwechsel in Betracht gezogen werden. Im CAT-System *Across* lassen sich beispielsweise über spezielle Einstellungen die internen Regeln zur Segmentierung anpassen. Für die Erkennung von

Zeilenwechselln müssten demnach im System entsprechende Bedingungen formuliert werden, die das Setzen unerwünschter Segmentgrenzen zuverlässig verhindern.

Falls die Änderung der programminternen Segmentierungsregeln im verwendeten CAT-System nicht möglich oder nicht gewünscht ist, so bleibt dem Übersetzer die Möglichkeit, den einzulesenden Text vor dem Alignment manuell nachzubearbeiten. Das bedeutet allerdings im Hinblick auf das dargestellte Problem, dass in Einzelfällen die getrennt stehenden Einheiten von Kollokationen in ein Segment künstlich zusammengeführt werden müssen. Dieser zusätzliche, jedoch einmalige Arbeitsschritt ist vergleichsweise aufwändig, dennoch erweist er sich als sinnvoll, da somit eine CAT-fähige Master-Fassung des englischen und deutschen Abkommens verfügbar wird. Aus einer solchen Vorlage lassen sich nachfolgend zahlreiche weitere Abkommen für andere englischsprachige Länder und Kulturen halbautomatisiert herleiten.

Im Alignment des deutsch-kanadischen Abkommens müssen dazu der ausgangs- und der zielsprachliche Text in jedem Fall manuell durchgesehen und die phraseologischen Einheiten in beiden Sprachen in die einander zugeordneten Segmente überführt werden. Anschließend sind im Translation Memory Satzpaare gespeichert, die für alle Fachtexte dieser Sprachkombination und damit für alle zu übersetzenden Abkommen zwischen Deutschland und englischsprachigen Ländern nutzbar sind. Die beschriebene Nachbearbeitung ist erst bei einer neuen Sprachkombination erneut erforderlich: Bei der Verabschiedung eines neuen DBA zwischen Deutschland und einem Land mit anderer Sprache (beispielsweise Spanisch) sind dann Ausgangstext und Zieltext vor dem Alignment wiederum entsprechend anzupassen.

Durch die manuelle Bearbeitung entstehen wieder Segmente, die sich in der Ausgangs- und in der Zielsprache inhaltlich entsprechen:

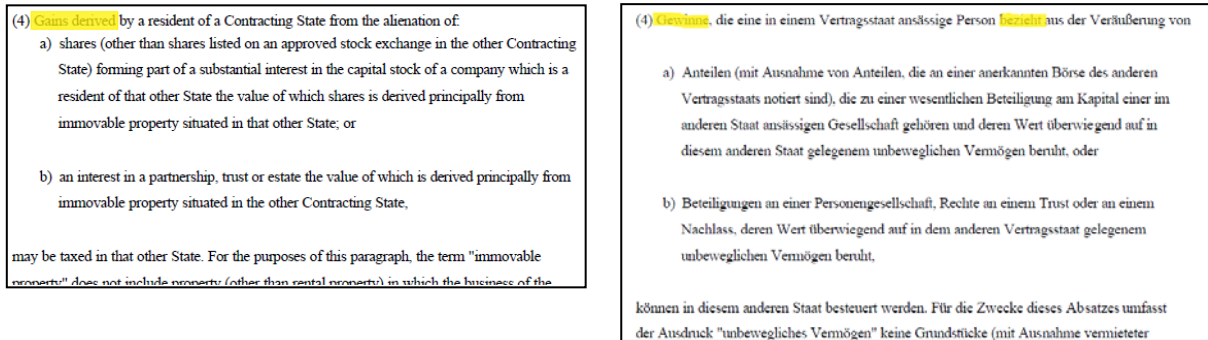


Abbildung 13: Nachbearbeitung und inhaltliche Entsprechung der Segmente

Unabhängig von der zuvor vorgeschlagenen Lösung erhält sich jedoch ein anderes Problem. Dieses resultiert aus den überlangen Sätzen und den überdehnten hypotaktischen Satzstrukturen, die für die Textsorte der Doppelbesteuerungsabkommen charakteristisch sind. Diese Satzstrukturen generieren wiederum große bis übergroße Segmente im Translation Memory. Zudem findet das TM-System Across im Übersetzungsspeicher nur solche Segmente, die zu mindestens 50 Prozent mit dem zu übersetzenden Segment übereinstimmen. Dies lässt sich in der nachfolgenden Abbildung 14 am Beispiel des DBA mit den Vereinigten Staaten von Amerika grafisch sichtbar machen:

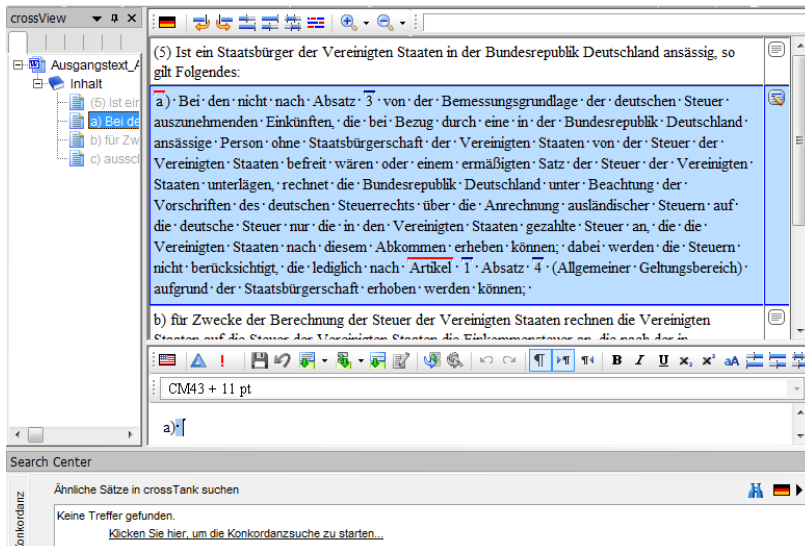


Abbildung 14: Hypotaktischer Satz als überlanges Segment

Die in Abbildung 14 dargestellte Problematik beschränkt sich nicht auf ein einzelnes DBA, sondern ist charakteristisch für die Texte der DBA zwischen Deutschland und anderen englischsprachigen Ländern. CAT-Systeme sind für hochfachsprachliche Texte mit einer

ausgeprägten Terminologie ausgelegt. Der Verarbeitung von Texten mit extremen hypotaktischen Satzstrukturen sind jedoch im System Grenzen gesetzt. In diesem Kontext sei erstens angemerkt, dass bei einem erhöhten Platzbedarf für die Darstellung des ausgangssprachlichen Textsegments weniger Platz zur Eingabe des Zieltextes auf dem Bildschirm verbleibt und die Benutzerfreundlichkeit abnimmt.

Zweitens gilt grundsätzlich, dass in überlangen Satzstrukturen die 50-Prozent-Marke zur Wiedererkennung gespeicherter Übersetzungseinheiten im Translation Memory seltener erreicht wird. Im Falle des CAT-Systems *Trados* liegt die Marke bei 30 Prozent. Bei der gegebenen Textstruktur reicht auch dieser Schwellenwert nicht sicher aus, um gespeicherte Einheiten zu erkennen. Die Wahrscheinlichkeit, einen Treffer unter den gespeicherten Satzpaaren zu erzielen, nimmt somit ab, je länger die gespeicherten Sätze sind. Um einen Treffer zu erzielen, müsste die zu suchende Formulierung im Translation Memory auch in kürzeren gespeicherten Segmenten vorkommen, damit eine ausreichend große Schnittmenge mit den gespeicherten Segmenten wahrscheinlicher wird. Dies kann aufgrund der ausgeprägt hypotaktischen Textstruktur dieser Textsorte jedoch nicht allgemein vorausgesetzt werden.

Hier bietet sich folgende methodische Lösung an:

Hilfreich wäre es, wenn eine häufig wiederkehrende Phrase aus den bestehenden Segmentpaaren extrahiert und angezeigt würde, auch wenn ihr Anteil am gesamten Segment immer noch unter 50 beziehungsweise 30 Prozent liegt. In Abbildung 15 ist beispielsweise die Nominalphrase „eine in einem Vertragsstaat ansässige Person“ aus dem Doppelbesteuerungsabkommen zwischen Deutschland und Kanada optisch unterlegt. Sie kommt in der deutschsprachigen Version des Textes insgesamt 14-mal vor und sollte deshalb im Zieltext im Sinne der terminologischen Konsistenz wiederkehrend gleich übersetzt werden. Allerdings erbringt die automatische Suche häufig keine Treffer, da diese Nominalphrase weniger als 50 bzw. 30 Prozent des Segments ausmacht:

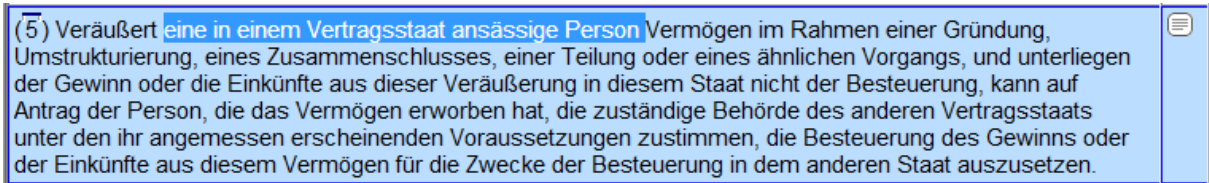


Abbildung 15: Nominalphrase in einem Segment als mögliches Strukturierungselement

Durch den geringen Anteil der Nominalphrase am gesamten Segment wird in der automatischen Suche von *Across* im Translation Memory *crossTank* kein Treffer angezeigt. Damit ist dieser Lösungsweg nicht durch die automatische Suche im Translation Memory zu erreichen:

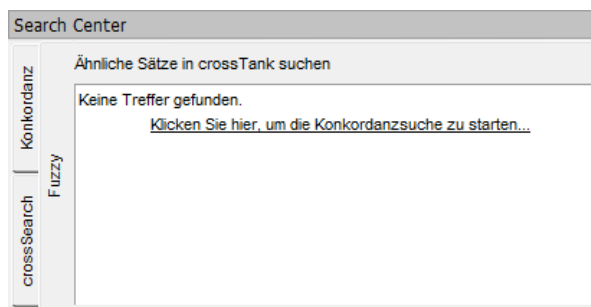


Abbildung 16: Die *crossTank*-Suche erfordert eine Mindestübereinstimmung von 50 Prozent. Theoretisch wären zwei Möglichkeiten denkbar. Zum einen könnte im Hintergrund eine Terminologiedatenbank gepflegt werden, die nicht nur Einwort- oder Mehrwortbenennungen als Termini, sondern auch Phrasen enthält. *Across* sucht im aktuell bearbeiteten Segment nach Wörtern oder Wortkombinationen, die als gespeicherte Termini in der Datenbank existieren, und zeigt diese im Fenster *crossTerm* daneben als Übersetzungsvorschläge an:

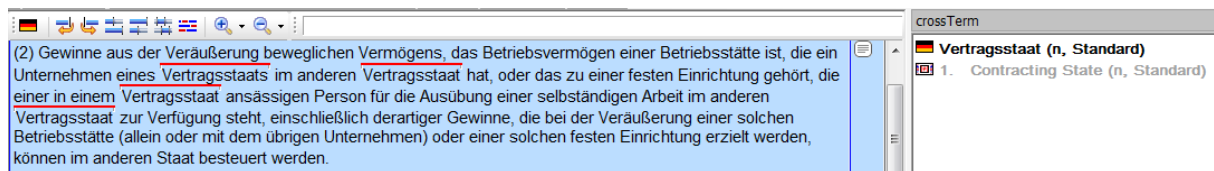


Abbildung 17: Die Terminologieerkennung in *Across*

Mit dieser vor allem für die terminologische Konsistenz nützlichen Funktion können auch kleine Ausschnitte eines Segments automatisch erkannt werden, die bei der Suche im Translation Memory nicht gefunden werden. Legt man anstelle des Lemmas „Vertragsstaat“



die gesamte Nominalphrase im terminologischen Eintrag an (ohne den unbestimmten Artikel), so wird diese bei der Terminologieerkennung wie ein einzelner Terminus im Segment gefunden:

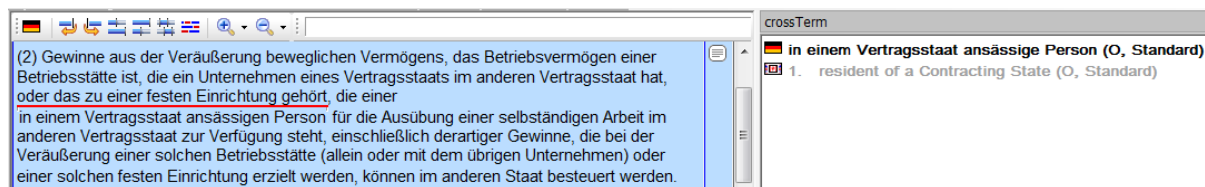


Abbildung 18: Die Terminologieerkennung ermöglicht das Finden kürzerer Segmentteile

Um die Terminologiedatenbank in der beschriebenen Weise zu nutzen, sind die entsprechenden Einträge vor dem Beginn der Übersetzung anzulegen. Als Termini nutzbare, häufig wiederkehrende Formulierungen kann der Übersetzer zudem erst bei einer Lektüre des Ausgangstextes identifizieren. Um die beschriebene Methode sinnvoll zu nutzen, bietet es sich aus diesen Gründen an, die Erfassung von Phrasen als terminologische Einträge mit der in Abbildung 13 gezeigten Nachbearbeitung des Ausgangstextes vor dem Alignment zu verbinden.

Alternativ wäre es möglich, im Alignment oder bereits im Ausgangstext die größeren Textsegmente fachsprachlicher Texte in kleinere Untersegmente und gegebenenfalls bis auf die Phrasenebene zu strukturieren, um so die Beschränkung durch den Mindestgrad der Übereinstimmung (50 beziehungsweise 30 Prozent) bei ausgangs- und zielsprachlichem Segment aufzuheben. Damit würde sich die Trefferwahrscheinlichkeit in *crossSearch*, der automatischen Suche innerhalb des Translation Memorys, erhöhen. Wäre die mittlere Länge eines Segments geringer, so hätte dies Auswirkungen auf die Berechnung des Übereinstimmungsgrades von dem gespeicherten und dem zu übersetzenden Segment. Eine geringere Segmentlänge würde somit zu tendenziell größeren Übereinstimmungen für gegebene Segmente führen.

Dieses vom Übersetzer durchzuführende Verfahren geht in der Praxis jedoch mit einem erheblichen Mehraufwand bei der Projektausführung einher. Selbst bei der Verwendung automatischer Prozesse wie dem Parsing durch Algorithmen obliegt dem Übersetzer anschließend eine zusätzliche Vorarbeit: Dazu gehört die Zuordnung der entstandenen

Phrasen in Ausgangs- und Zielsprache, bevor oder während die zu alignierenden Dokumente in das TM-System eingelesen werden. Zusätzlich ist bei automatischen Systemen eine maximale Tiefe des Parsings im Syntaxbaum festzulegen und es bleibt offen, ob für jede Sprache ohne weitere Nacharbeit ein verwendbares Ergebnis entsteht.

## 5 Die Konkordanzsuche im Translation Memory und die Einbindung einer Terminologiedatenbank

### 5.1 Die Konkordanzsuche

In einem weiteren Verfahrensschritt wird nun geprüft, ob sich eine Problemlösung über die im Programm implementierte und angebotene Konkordanzsuche im Translation Memory erzielen lässt. Zusätzlich werden die Möglichkeiten der in der Praxis gepflegten terminologischen Eintragsstruktur in der Terminologiedatenbank betrachtet.

Führt man die Konkordanzsuche durch, so erweist sie sich dann als besonders effektiv, wenn vergleichsweise kurze Phrasen oder Schlagworte zum Gegenstand der Suche werden. Die Suche verläuft jedoch nicht erfolgreich, wenn sich der Suchausdruck in Form eines Teilsatzes, einer Phrase, eines Mehrwortterminus oder eines Phraseologismus über mehrere Segmente erstreckt. In diesem Fall werden beide davon betroffenen Segmente nicht erkannt. Diese Situation kann auftreten, wenn die in Abbildung 11 gezeigten unerwünschten Zeilenwechsel nicht mit einer der vorgestellten Methoden neutralisiert wurden. Ein entsprechendes Beispiel ist nachstehend abgebildet: Die fachsprachliche Kollokation „Veräußerung von Anteilen“ wird von der Konkordanzsuche nicht gefunden.

(3) Gewinne eines Unternehmens aus der Veräußerung von Seeschiffen, Luftfahrzeugen und Containern, die im internationalen Verkehr betrieben beziehungsweise genutzt werden, oder aus beweglichem Vermögen, das dem Betrieb dieser Schiffe oder Luftfahrzeuge dient, können nur in diesem Staat besteuert werden.
(4) Gewinne, die eine in einem Vertragsstaat ansässige Person bezieht aus der Veräußerung von
a) Anteilen (mit Ausnahme von Anteilen, die an einer anerkannten Börse des anderen Vertragsstaats notiert sind), die zu einer wesentlichen Beteiligung am Kapital einer im anderen Staat ansässigen Gesellschaft gehören und deren Wert überwiegend auf in diesem anderen Staat gelegenes unbewegliches Vermögen beruht, oder
b) Beteiligungen an einer Personengesellschaft, Rechte an einem Trust oder an einem Nachlass, deren Wert überwiegend auf in dem anderen Vertragsstaat gelegenes unbewegliches Vermögen beruht,

Abbildung 19: Suche über Segmentgrenzen hinweg

Um diese Situation zu verhindern, bietet es sich an, nur kurze Sucheinheiten, die wahrscheinlich nicht durch die Textgestaltung in der Ausgangssprache voneinander getrennt werden, für die Konkordanzsuche zu verwenden.

In der Praxis ergibt sich eine Einschränkung dadurch, dass in CAT-Systemen innerhalb des zielsprachlichen Segments die gesuchte Phrase oder der Terminus optisch nicht hinterlegt wird. Das ist jedoch kein grundsätzlicher Nachteil. Auf der Basis seiner fremdsprachlichen Kompetenz ist der Übersetzer in der Lage, die zielsprachliche Einheit im Segment zu identifizieren und gegebenenfalls in seine Übersetzung zu integrieren. Es ist für die Konkordanzsuche somit nicht erforderlich, sich auf einzelne Wörter als Suchausdrücke zu beschränken; auch kurze Phrasen oder Kollokationen können auf diesem Weg gefunden werden. Ergibt eine solche Suche keinen Treffer, so existiert entweder keine Fundstelle oder der gesuchte Ausdruck ist durch die Textformatierung typografisch getrennt. Im letzteren Fall kann sich der Übersetzer durch die Verwendung eines kürzeren Suchausdrucks vergewissern, dass die gesuchte Phrase tatsächlich nicht vorkommt.

## 5.2 Die Terminologiedatenbank

Ein zusätzliches Problem verbindet sich damit, dass terminologische Einträge in Fachdatenbanken oftmals nicht lexikografisch ausgearbeitet sind. Dies gilt vor allem im Hinblick auf Fachkollokationen, da die Erschließung dieses phraseologischen Materials mit einem erheblichen Zeit- und Kostenaufwand verbunden ist. Im Regelfall ist bei der Suche nach einzelnen Termini eine eingeschränkte Terminologiedatenbank im Hintergrund verfügbar. Diese durchsucht das aktuelle ausgangssprachliche Segment nach dem Terminus und zeigt das gespeicherte zielsprachliche Äquivalent als Vorschlag an. Allerdings hat die Suche über eine Terminologieverwaltung in der Praxis eben den Nachteil, dass sie häufig über die zielsprachliche Benennung hinaus keine Informationen zur Textproduktion bereitstellt. Dies erklärt sich daraus, dass in der terminologischen Praxis keine umfassende Eintragsstruktur gepflegt wird, in der auch Kontextbeispiele, Kollokationen und freie Kombinationen in beiden Sprachen hinterlegt werden. Selbst wenn eine solch umfangreiche Eintragsstruktur bei der Terminologiearbeit gepflegt wird, wäre es außerdem erforderlich,

im Sichtfenster des Terminologiemoduls (vgl. Abbildung 17) den gefundenen Terminus auszuwählen und zum Aufrufen der Details in ein separates Fenster zu wechseln. Dies unterbricht den Arbeitsprozess und relativiert den Nutzen von umfangreich gepflegten, phraseologisch erweiterten Terminologieeinträgen. Die in Abbildung 18 gezeigte Nutzung einer Phrase als Lemma ermöglicht dagegen zumindest eine eingeschränkte Erweiterung der angebotenen terminologischen Informationen, ohne die Arbeitsumgebung verlassen zu müssen.

Mit der Konkordanzsuche im Translation Memory bietet sich hier eine alternative Problemlösung, da die Treffer im Kontext angezeigt werden. Im besten Fall enthält eines der Kontextbeispiele außer dem zielsprachlichen Äquivalent ebenfalls einen vom Übersetzer gesuchten Kollokator in der Zielsprache. Die Wahrscheinlichkeit dieser für die Übersetzung günstigen Konstellation erhöht sich, wenn zuvor entsprechend der Abbildung 13 eine durch die Segmentierung getrennte Kollokation per manueller Nachbearbeitung wieder zusammengeführt wurde. Dieses modifizierte Segment wird dann in der Konkordanzsuche so gefunden, dass die Kollokation in beiden Sprachen vollständig enthalten ist. Im nachfolgenden Beispiel erhält man zum Terminus „Gewinne“ neben der Übersetzung „gains“ gleichzeitig den Kollokator „beziehen“ mit der Übersetzung „to derive“:

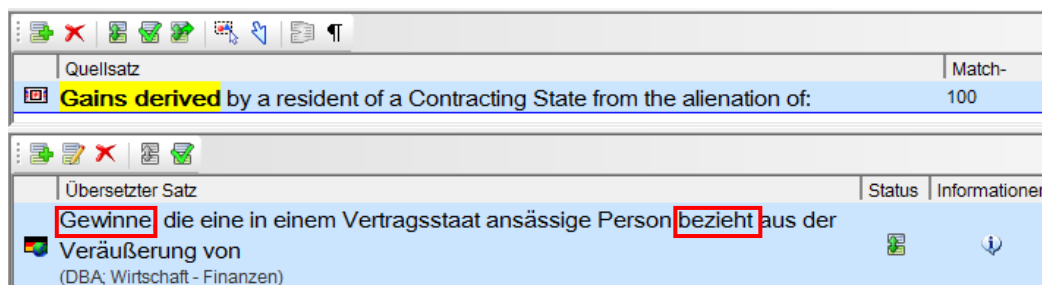


Abbildung 20: Fachphraseologische Zusatzinformationen über die Konkordanzsuche

Aus diesem Grund ist bis zu einem gewissen Grad die in Kapitel 4 erwähnte phraseologisch motivierte Bearbeitung der Segmente sinnvoll, selbst wenn keine phrasenbasierte Segmentierung per Parsing erstellt wurde. Die so gewonnenen fachsprachlichen Kollokationen sind textsortenauthentisch und können für die gegebene Übersetzung als gesichert gelten.

Dieser Weg ist der Qualitätssicherung förderlicher als eine zeitaufwändige Recherche zielsprachlicher Formulierungen im Internet, die oftmals zu nicht verifizierten Quellen führt. Im Fall der aus länderspezifischen DBA erzeugten TM-Daten ergibt sich hier auch der Vorteil, dass die offiziellen Benennungen der deutschen Steuern wie bei Terminologieeinträgen im Translation Memory gefunden werden, vergleiche hierzu Abbildung 21:

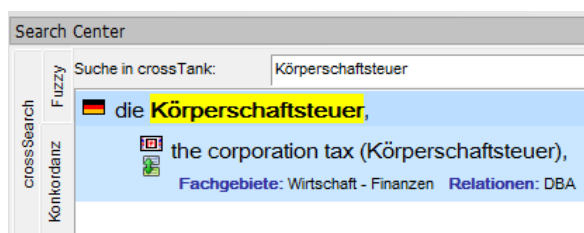


Abbildung 21: Terminologische Informationen über die Konkordanzsuche

Im Folgenden soll gezeigt werden, dass die integrierte Konkordanzsuche im CAT-System dafür eingesetzt werden kann, innerhalb gespeicherter Segmentpaare zielsprachliche Termini und Kollokationen des Ausgangstextes zu identifizieren. Dabei wird sich auch zeigen, dass der zunächst eingeschlagene methodische Weg der phrasenbasierten Segmentierung wesentlich zeitintensiver ist als die Nutzung der Konkordanzsuche, weshalb die Autoren dieses Artikels diesen Weg bevorzugen.

Insbesondere bei Fachtexten mit durchgehend großen Segmentlängen erweist sich eine manuelle Konkordanzsuche als erfolgversprechendes methodisches Hilfsmittel. Für diese Konkordanzsuche ist im Gegensatz zu der automatischen Suche innerhalb des Translation Memorys eine größere Segmentlänge förderlich. Je größer ein zu durchsuchendes Segment ist, desto wahrscheinlicher liegt ein Suchausdruck innerhalb dieses Segments. Damit tritt der zuvor beschriebene Fall seltener ein, dass ein Suchausdruck wegen der typographischen Zerteilung vom CAT-System nicht erkannt wird. Dies bietet im Vergleich zu der zuvor genannten Methode der Fragmentierung von Segmenten den Vorteil, dass mit den hypotaktischen Strukturen der Textsorte unverändert gearbeitet werden kann. Damit entfällt ein manuelles Aufteilen der Segmente in Phrasen, das einen vergleichsweise aufwändigen Verfahrensschritt darstellt.

Jedoch ist auch bei der manuellen Konkordanzsuche, die als Lösungsweg vorgeschlagen wurde, eine manuelle Handlung des Übersetzers erforderlich. Die Konkordanzsuche muss nach dem Markieren des Suchausdrucks über ein Kontextmenü oder per Tastenkombination ausgelöst werden, da das Korpus vom Programm nicht automatisch durchsucht wird.

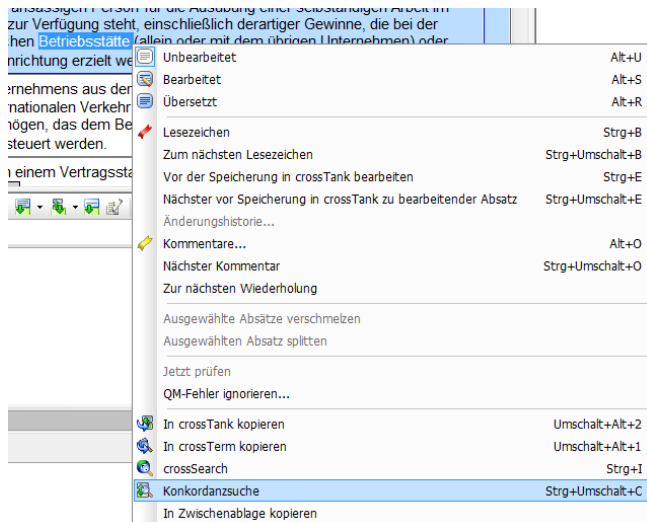


Abbildung 22: Die manuelle Konkordanzsuche über das Kontextmenü

## 6 Fazit

Mit dieser Nutzung der Konkordanzsuche verbinden sich mehrere erkennbare methodische Fortschritte. Zunächst ist eine zeitaufwändige übersetzungsbezogene Feindifferenzierung der Segmente nach Phrasen nicht mehr notwendig. Die Konkordanzsuche stellt Satzbeispiele bereit, die einen phraseologischen Mehrwert liefern. Diese Art der Suche ist einfach, erspart Zeit und lässt sich individuell an den benötigten Problemstellen des Textes einsetzen.

Darüber hinaus ist die Konkordanzsuche vor allem auf die gegebenen, überlangen Segmente der Textsorte hin ausgerichtet.

Die Erfahrungen belegen, dass sich bei Übersetzungen von steuerrechtlichen Fachtexten mit einem CAT-System die integrierte Konkordanzsuche als Mittel der Qualitätssicherung und als Ressource für fachsprachliche Phraseologie bewährt. Die von *Across* bereitgestellte Konkordanzsuche in TM-Einheiten erweist sich als wirksames Arbeitsmittel, um fachphraseologische Einheiten in der Zielsprache zu finden.

## 7 Bibliographie

Abkommen zwischen der Bundesrepublik Deutschland und dem Vereinigten Königreich Großbritannien und Nordirland zur Vermeidung der Doppelbesteuerung und zur Verhinderung der Steuerverkürzung auf dem Gebiet der Steuern vom Einkommen und vom Vermögen. *Bundesgesetzblatt Jahrgang 2010 Teil II Nr. 33, ausgegeben zu Bonn am 23. November 2010, S. 1334ff.*

Abkommen zwischen der Bundesrepublik Deutschland und den Vereinigten Staaten von Amerika zur Vermeidung der Doppelbesteuerung und zur Verhinderung der Steuerverkürzung auf dem Gebiet der Steuern vom Einkommen und vom Vermögen und einiger anderer Steuern. *Bundesgesetzblatt Jahrgang 2008 Teil II Nr. 15, ausgegeben zu Bonn am 23. Juni 2008, S. 612ff.*

Abkommen zwischen der Bundesrepublik Deutschland und der Republik Singapur zur Vermeidung der Doppelbesteuerung auf dem Gebiet der Steuern vom Einkommen und vom Vermögen. *Bundesgesetzblatt Jahrgang 2006 Teil II Nr. 27, ausgegeben zu Bonn am 3. November 2006, S. 931ff.*

Abkommen zwischen der Bundesrepublik Deutschland und Kanada zur Vermeidung der Doppelbesteuerung auf dem Gebiet der Steuern vom Einkommen und bestimmter anderer Steuern, zur Verhinderung der Steuerverkürzung und zur Amtshilfe in Steuersachen. *Bundesgesetzblatt Jahrgang 2002 Teil II Nr. 11, ausgegeben zu Bonn am 27. März 2002, S. 671ff.*

Across - Sprachtechnologie für effizientes Übersetzungsmanagement. <http://www.across.net>.

Agreement between the Federal Republic of Germany and Canada for the Avoidance of Double Taxation with respect to Taxes on Income and Certain other Taxes, the Prevention of Fiscal Evasion and the Assistance in Tax Matters. *Bundesgesetzblatt Jahrgang 2002 Teil II Nr. 11, ausgegeben zu Bonn am 27. März 2002, S. 671ff.*

Agreement between the Federal Republic of Germany and the Republic of Singapore for the Avoidance of Double Taxation with respect to Taxes on Income and on Capital.

*Bundesgesetzblatt Jahrgang 2006 Teil II Nr. 27, ausgegeben zu Bonn am 3. November 2006, S. 931ff.*

Brähler, Gernot (2012). *Internationales Steuerrecht. Grundlagen für Studium und Steuerberaterprüfung*. 7. Aufl. Wiesbaden: Springer Gabler.

Convention between the Federal Republic of Germany and the United Kingdom of Great Britain and Northern Ireland for the Avoidance of Double Taxation and the Prevention of Fiscal Evasion with Respect to Taxes on Income and on Capital. *Bundesgesetzblatt Jahrgang 2010 Teil II Nr. 33, ausgegeben zu Bonn am 23. November 2010, S. 1334ff.*

Convention between the Federal Republic of Germany and the United States of America for the Avoidance of Double Taxation and the Prevention of Fiscal Evasion with respect to Taxes on Income and Capital and to certain other Taxes. *Bundesgesetzblatt Jahrgang 2008 Teil II Nr. 15, ausgegeben zu Bonn am 23. Juni 2008, S. 612ff.*

Haase, Florian (2011). *Internationales und Europäisches Steuerrecht*, 3. Aufl. Heidelberg: C.F. Müller.

Kramer, Jörg-Dietrich (1990). *Grundzüge des US-amerikanischen Steuerrechts*. Stuttgart: Poeschel.

Organisation for Economic Co-operation and Development (OECD) (2010). *Model Tax Convention on Income and on Capital. Condensed Version. 22 July 2010.*

Tipke, Klaus und Joachim Lang (2013). *Steuerrecht*. 21., völlig überarbeitete Aufl. Köln: Verlag Dr. Otto Schmidt.

Vogel, Klaus und Moris Lehner (2008). *Doppelbesteuerungsabkommen: DBA der Bundesrepublik Deutschland auf dem Gebiet der Steuern vom Einkommen und Vermögen. Kommentar auf der Grundlage der Musterabkommen*. 5., völlig neubearbeitete Aufl. München: C.H. Beck.



## **T21N - Translation in Transition**

T21N offers a cutting-edge electronic publishing venue, created by experts for both young talent and established researchers from the worlds of translation and interpreting.

T21N provides a stage for emerging ideas and new academic talent to present their ideas in a digital reading site, where speed and ease meet enjoyment.

T21N is exclusively published online at <http://www.t21n.com>.

Articles in compliance with our style sheet may be submitted at any time and will be published at short notice.

T21N editors research and teach at the Institute of Translation and Interpreting at the University of Heidelberg in Germany.

Editors:

Dipl.-Übers. Viktorija Bilić, Dr. Anja Holderbaum,  
Dr. Anne Kimmes, Prof. Dr. Joachim Kornelius,  
Dr. John Stewart, Dr. Christoph Stoll