

# LANGUAGE TECH NEWS

*ata* A member of the American Translators Association.

## Editor's Note:

Here we are, already well into the year 2008 and now in Volume 2 of our newsletter. Again, huge thanks go to all of those who contributed. I hope to continue to gather articles, reviews, and news from old and new contributors alike to keep this a vibrant, useful tool for division members.

Many of us are already fully occupied (overloaded?!) with work and home responsibilities, and don't always have time to submit an article. However, if you happen to be surfing or reading and find industry-related articles of particular interest, let me know. I can always try to get reprint permission to share it with your colleagues in an upcoming issue.

For now, enjoy the offerings here! □

*Lisa Carter*

*Editor*

**VOL. 2, NO. 1 / MARCH 2008**

### IN THIS ISSUE:

Assistant Administrator's Letter .....	2
Quick Tips .....	3
Turbocharging Your QA: Regular Expressions for Translators .....	8
A Slightly Faster and Safer Internet .....	3
Trados Tip .....	11

### Call for Reviewers

Are you using software that helps you in your day-to-day work as a translator and/or interpreter? Tell us about it!

Send reviews of your new, favorite, even most-hated language technology software to [lisa@intralingo.com](mailto:lisa@intralingo.com) by the submission deadlines.

Your colleagues will thank you!

# Assistant Administrator's Letter

Welcome to this issue of the LTD Newsletter and welcome to the LTD if you are a new member. As assistant administrator, my job is to keep in touch with the membership and aid in the exchange of information. According to our [bylaws](#), some of the principal objectives of the division are:

- To provide information to its members in areas related to language technologies,
- To further cooperation and information exchange among its members, and
- To serve as a forum for discussion through its publications and other venues.

In addition to the [LTD mailing list](#), Michael Wahlster (the webmaster) and I have been contributing regularly to the [LTD blog](#) with information on tools and events, and I have been creating pages on specific categories of tools on the [LTD site](#), such as [Text-to-Speech Tools](#). We look forward to your comments on the blog, on the information provided on the site, and to your replies on the mailing list. Each one of us knows something that many other members do not, so share your knowledge! And don't forget about the newsletter, which awaits your contributions. □

*Sincerely,*

*Naomi J. Sutcliffe de Moraes*

*Assistant Administrator*

Language Tech News

Vol. 2, No. 1 March 2008

Copyright©2008  
American Translators  
Association

**Editor**

Lisa Carter

[lisa@intralingo.com](mailto:lisa@intralingo.com)

**Proofreaders:**

Barbara Guggemos

Naomi J. Sutcliffe de Moraes

**Contributors to this Issue:**

Carey Holzman

Tuomas Kostianen

Jonathan Lukens

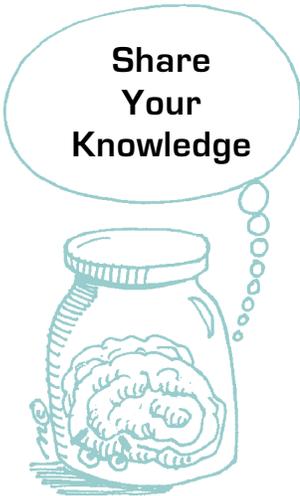
Naomi J. Sutcliffe de Moraes

**Layout:**

Cindy Gresham

[illustrator@cgrisham.com](mailto:illustrator@cgrisham.com)

LTD is the  
Language Technology Division  
of the American Translators  
Association



Share  
Your  
Knowledge

*Language Tech News is a great way to get your name out to your colleagues, to share your expertise, and give back to your division, your association. Send in an article and share the wealth of knowledge you have!*

## Keep Up-to-Date

Log on to the division website where you'll find current information about the LTD and related topics, including:

- CAT Tools
- Dates and Events
- Linux
- Mac OS
- Windows
- Worth A Look
- LTD Contacts

[www.ata-divisions.org/LTD](http://www.ata-divisions.org/LTD)

# Quick Tips

by Naomi J. Sutcliffe de Moraes

## Tip #1: DéjàVuX Tags

I recently had a discussion with a colleague who has been using DéjàVu (a translation environment tool) for years. He complained about tags that are not needed in the translation. For example, it is not that uncommon for a foreign word to be italicized in an original text. When the text is translated into that language, the italics are often no longer needed. In Figure 1, codes {1} and {4} are for formatting applied to the entire sentence and codes {2} and {3} are for the italics applied to the word “italics.” These codes are actually how MS Word represents italics internally (sometimes I miss WordPerfect, which would let you see the codes and move them around in a separate window). **However, there is nothing wrong with putting the two italics codes next to each other, effectively cancelling each other out.** This is shown in the right part

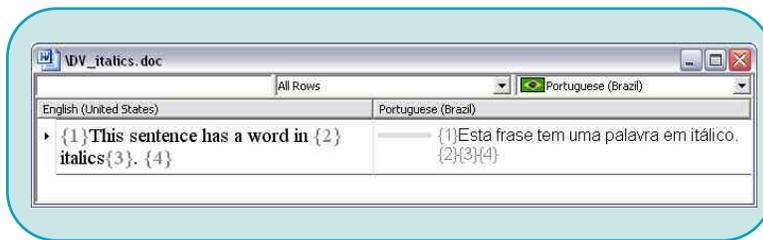


Figure 1 – DéjàVuX editing screen showing codes.

of the window.

Sometimes, when I convert a file from PDF, I end up with a lot of junk formatting—sometimes with codes between every word! I keep an eye on the original text in a separate window and:

- Put any codes appearing at the beginning of the sentence at the beginning in the translation, and
- Put any codes appearing in the body of the text at the end, IN THE SAME ORDER.

This has worked wonderfully, though reading converted text with all those codes is a real pain in the rear.

Note that DéjàVuX allows you to **edit the original text**, the only tool I have used that does. You can even delete codes. This is best avoided, if possible, but every once in a while I will get something like 1°. Sometimes the ° is a character, other times it is a code. Since this is 1<sup>st</sup> in English, I don't need the °. In this case, the code

can be deleted in the original (left pane) and then not used in the translation (right pane).

## Tip #2: DéjàVuX Key Combinations

In my long review of DéjàVuX in the August and September issues of *The ATA Chronicle*, I mentioned my frustration at not being able to change the case of a word pulled in from the terminology database. Basically, if I needed “corporation” and the match from the database was “Corporation,” the shortcut CTRL+3 (or whatever the corresponding number was) would insert the word with an initial capital letter. I am now happy to admit I was wrong! Typing SHIFT+CTRL+3 will insert “corporation” with the lowercase “c.” Another nice feature which I have only just begun to use is CTRL+RIGHT ARROW to move to the right one word at a time and likewise for the left arrow. This is great when you need to jump over a few words, but not all the way to the end of the line. This also works in MS Word.

## Tip #3: Getting to Know Microsoft Word

As MS Word users, we are often content with knowing the minimum we need to write what we want. However, clients can send documents that incorporate features we have no idea how to use. For example, if you do not know how to create a table of contents (TOC) automatically using styles, you should read up on it because, if not, you could change the codes creating the TOC without meaning to and then an entry will be missing. Typing a TOC, even if the original document is provided scanned, is a quality-assurance nightmare. Any change in formatting will change all the page numbers and, when changing a title in the text, you must remember to change the title in the TOC. Learn all you can about styles, because they can make your life much easier.

I love manuals (naturally, since I used to write them) and am very happy with my *Special Edition Using Microsoft Office Word 2003* by Bill Camarda. I recently installed Word 2007 and love it—so much that I uninstalled Word 2003 and must now buy a new manual. I like that Camarda's book deals both with the basics for beginners and with the really complicated things I

sometimes need. I also like that it honestly tells you when a feature is buggy and should be avoided. The Word 2007 book by the same publisher is written by Faithe Wempen, and I hope it is just as good.

#### Tip #4: Layout in Microsoft Word with Tables

Tables are my personal favorite Word feature. They are great for laying out text to match an original not provided in MS Word. Most of us were taught to use tabs to create space on a page back in typewriter days, but this is not the best option for translation. Some documents I receive contain five spaces rather than a tab, which is even worse because spaces can change size based on justification. These are the advantages of using tables:

- You can hide the table lines/borders any way you like, and the reader will not necessarily notice the text is in a table.
- If you (or the client) later change the font, font size or page layout, the text will wrap around within the table rather than pushing the following text where it shouldn't go.

- You can justify the text in each cell in a different way to get different effects. You can even put an image in one cell and text in the other, and their relative locations will be completely fixed.

Programs that convert PDF files into text files often lay text out using tabs. This is a disaster when the target language occupies significantly more or less space than the source language. I prefer to preprocess (reformat) files converted by PDF conversion programs in order to make my life easier later.

Turn tables created with tabs into proper tables (in MS Word 2007) by selecting the text. Then, in the **Insert** menu, click on **Table** and then **Convert Text into Table**. This was done in Figure 2.

I find tables great for diplomas or official documents with lots of stamps and seals. Headers also often need this type of formatting. Figure 3 shows two examples. The table borders have been drawn in (red) but do not appear in the document. □

**Naomi J. Sutcliffe de Moraes** is a Portuguese into English translator. She has a PhD in Linguistics from the University of São Paulo, but prior to becoming a translator she worked as an engineer and physicist. Her specialties are medicine, engineering, and law. She can be reached through [www.JustRightCommunications.com](http://www.JustRightCommunications.com)

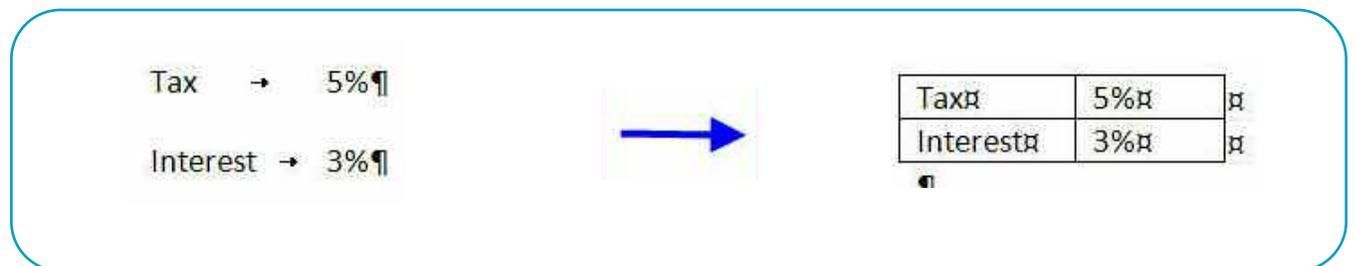


Figure 2, above:  
Converting text with tabs into a table.

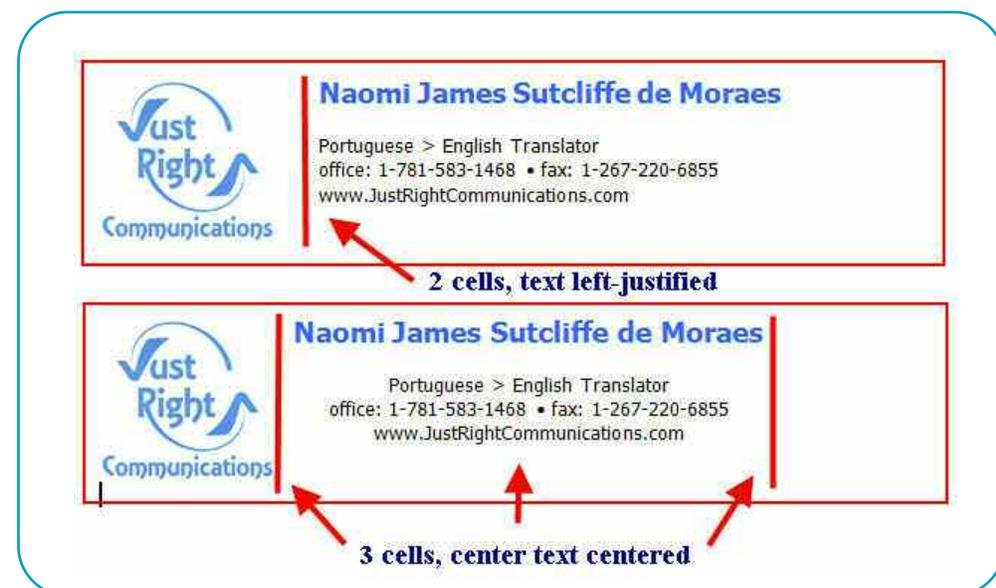


Figure 3, left:  
Use of tables to duplicate difficult formatting in original document.

# Turbocharging Your QA: Regular Expressions for Translators

by Jonathan Lukens

## It's 4:30 p.m. You are two paragraphs from the end of a

two-week translation job, due the following morning, when a chime from your e-mail alerts you to a new message from your client:

“You are using European format for the dates, right?”

You check the purchase order and all of your correspondence: this is the first you have heard of this request. You find this especially frustrating because you have already made dinner plans for the evening. Your options are: (a) cancel all engagements and plod through the project's 35,000 words looking for dates to correct; (b) tell the client to handle the problem themselves since you've done everything that was asked of you; or (c) replace ([ 0-1 ] ? [ 0-9 ] ) / ( [ 0-3 ] ? [ 0-9 ] ) / ( ( [ 1-2 ] [ 0-9 ] ) ? [ 0-9 ] { 2 } ) with \$ 2 / \$ 1 / \$ 3.

The winning answer is (c). This is not a monstrous typo, but rather a regular expression. A regular expression, often abbreviated to “regex,” is a way of telling a computer how to find a pattern in a text.

Most computer users know how to use the star wildcard (\*). For example, if you want to search for all of the PDF files on your computer, you could search for \*.pdf. Regexes enable you to be much more specific. If you only want to search for PDFs whose titles begin with two letters followed by a sequence of any number of digits, you can. You can find all of the instances of a word that has multiple spellings, such as grey and gray, or Dmitri, Dmitriy, and Dmitrij. You can, as hinted above, replace January 10, 2007 with 10 January 2007 without actually specifying the month, day, or year. Regular expressions can turn a two-hour QA headache into a five-minute snap. This means spending less time to produce a higher-quality finished product.

Although regular expressions originated in the programming world, their tremendous utility has led software developers to integrate them into many common programs, such as Dreamweaver and Microsoft Word. Using regexes can be as easy as pressing Ctrl+F in these applications. Dreamweaver is probably the best option for trying out the scenarios presented in this introduction to using regexes. (You can download a 30-day trial of Dreamweaver from the Adobe website at [www.adobe.com](http://www.adobe.com).) The example given in option (c) above will be used for illustration.

### Comparison of Dreamweaver and MS Word Regular Expression Syntax

Dreamweaver	Microsoft Word
[ ^ ]	[ ! ]
.	?
?	n/a
*	n/a
{ }	Supported, except that you cannot assert zero instances of something using { 0, } or { 0, n }
	n/a
\ b	< (beginning of word), > (end of word)
\$ n	\ n
^	n/a
\$	^ 1 3

### Getting Started with Regular Expressions

To search a file using regexes in Dreamweaver, press Ctrl+F and select the “Use regular expressions” checkbox. You are now ready to use regular expressions. Setup is

equally simple in Microsoft Word—just open the Find window, click “More >>” and select the “Use wildcards” checkbox. However, the Word version of regexes is far less powerful than the standard regexes used in Dreamweaver; hence this article’s use of Dreamweaver syntax. The reference table included on the previous page shows the distinctions and limitations of Word regexes.

Regular expressions are, in a certain sense, a language; therefore, your aim should be fluency. The faster you can translate a pattern in English into a jumble of symbols like the example above, the better they will serve you. As you learn the concepts of regex syntax, it will help to read examples out loud, translating them into ordinary English. The better you can read regexes, the faster you will be able to write them.

For clarity, all regex text in this article is in Courier New, and all searched text is **boldface**.

### Character Classes

Regexes use a set of square brackets to group all the potential matches for a single character. This group is referred to as a *character class*. If you want to find “a,” “b,” or “c,” you create a character class thus: `[abc]`. “Any vowel” translates as `[aeiou]` or “any even numeral” would be `[02468]`.

Let’s return to the example we began with and find the first character class: `[0-3]`. The hyphen between the two characters indicates a range of characters to match. Translated into English, this reads as “any numeral from zero through three.” This applies to letters as well: `[a-z]` matches any lowercase Latin letter and `[A-Z]` matches any uppercase Latin letter. You can combine these series to fit your needs. “Any upper- or lowercase Latin letter” will be `[a-zA-Z]`, and `[a-mo-z]` will match every lowercase Latin letter except “n”.

What if you want to say “Anything but xyz”? You can negate a character class with a caret. Thus, `[^0-9]` will find any character that is not a numeral and `[^aeiou]` will skip all vowels.

(Note: if you want to include a literal hyphen in a character class, it must go at the very end. The regex `[0-9-]` would match any keystroke in a hyphenated phone number. Outside a character class, a hyphen is just a plain old hyphen. Carets also only negate within a character class.)

A dot, or period (`.`), is a shortcut meaning all characters, including upper- and lowercase letters, punctuation, and numbers. It is a self-standing character class and therefore does not need the usual square brackets. `C.PO` will therefore match both **CAPO** and **C3PO**.

**Puzzle:** which of the options below will the following regular expression not match? Try reading it aloud. (Answers to puzzles are found in the corresponding footnotes.)<sup>1</sup>

`[^aeiou][aeiou][^aeiou]`.

- a) cord
- b) wow!
- c) arts
- d) win9

### Quantifiers and Parentheses

By default, character classes only match one character from the class at a time. Therefore, to search for any sequence of four digits, you would search for `[0-9][0-9][0-9][0-9]`. This is certainly faster than searching for every possible combination of four digits but somewhat clumsy. Fortunately, the special wildcards `?`, `*`, `+`, and `{ }` provide you with shortcuts to answering the question, “How many?”

A question mark translates as “zero or one of” the preceding characters. Therefore, searching for `runs?` will find both **run** and **runs**. Question marks come in handy when you are looking for something that has an optional element. Asterisks, usually referred to as “stars” in regex-speak, are read “zero or more of.” Therefore, `10*` will match **1**, **1000**, and **1000000**. Plus signs translate to “at least one.” This means that searching for `travel+ed` will match **traveled** and **travelled**. It will also match **travelled**, which may be more than you are looking for.

Curly brackets, used for finding a specific quantity of characters, are a little pickier. These come in three “flavors”: `{x}`, `{x,}`, and `{x, y}`. We only see one variety in our example. `[0-9]{2}` is a shortcut to `[0-9][0-9]`. `{x}` specifies an exact number of the preceding character. `{x,}` specifies at least `x` number of the preceding character, so `[a-zA-Z]{3,}` will match all words that are three letters long or longer. `{x, y}`

*Regular expressions can turn a two-hour QA headache into a five-minute snap. This means spending less time to produce a higher-quality finished product.*

will match any number of the preceding character from  $x$  to  $y$ , as in `[ 0-9 ] { 3, 5 }`, which will match all sequences of three, four, or five digits.

Parentheses provide another important regex feature. They allow you to group a given number of characters and apply any of the above quantifiers to them as one unit. Thus, where our example says `( [ 1-2] [ 0-9 ] ) ? [ 0-9 ] { 2 }`, it means that the pattern will match either **2006** or **06**, or either **1883** or **83**.

**Puzzle:** what would I be searching for with the following regex (only applicable to the U.S.):<sup>2</sup>

```
[ 0-9 ] { 5 } ( - [ 0-9 ] { 4 } ) ?
```

## Capturing and Substitution

Parentheses also serve another important function: they allow you to *capture* the text that you match. In our example, we searched for dates in a North American format. Then we replaced them with a mysterious `$ 2 / $ 1 / $ 3`. Each number, when combined with the preceding dollar sign, represents the contents of a set of parentheses. `$ 1` corresponds to the first set of parentheses, `$ 2` corresponds to the second set of parentheses, and `$ 3` corresponds to the third. As demonstrated, you can thus reuse what you find when replacing matches and even rearrange the order of what you capture.

**Puzzle:** You have a long list of names in a table. The names do not need translating. However, they are all given with the last name, followed by a comma, then the first name. You write the following regex:

```
( [ A-Z] [ a-z ] + ) , ( [ A-Z] [ a-z ] + )
```

What should you put in the “Replace” field to reverse the first and last names?<sup>3</sup>

## Alternation, Word Boundaries, and Escape Sequences

Another important regex operator is the vertical bar (`|`), which you may read as “or.” Using the bar, you can provide several alternatives when searching a text. The bar enables you to search for either “futures” or “stock” simultaneously with `futures | stock`.

Let’s say you have just completed a translation that deals with regulations on over-the-counter markets and the pharmaceutical industry. Perhaps you, like me, dictate all of your translations and your speech

recognition software never abbreviated **over-the-counter** to **OTC**. You want to abbreviate **over-the-counter** to **OTC** if it is followed by **market**, **contract**, or **derivatives**; in all other cases, **over-the-counter** refers to **over-the-counter drugs**. Search for `over-the-counter ( market | contract | derivatives )` and replace it with `OTC $ 1`. Problem solved.

What if an acronym such as “OTC” is not consistently capitalized? You could simply find and replace `otc` with `OTC`. However, this may yield unwanted results, like **ScOTCh**, **dOTCom**, and **top-nOTCh**. You could fix this by inserting a space after the search term, but you will then skip any instances of the acronym that are directly followed by a punctuation mark. Regexes provide you with a useful tool for working around this situation—they allow you to include word boundaries (`\ b`) in your searches. Thus, you can solve this problem by searching for `\ botc \ b`. The `\ b` translates into English as “begins with” or “ends with,” respectively.

Did that backslash in front of the `b` catch your attention? A regex engine interprets a backslash as a signal that the next character has a meaning other than what it normally has. The preceding backslash can give a character a meaning other than its literal meaning, as with `\ b`. It can also give a literal meaning to a character that the search would normally interpret symbolically. Therefore, to search for `?`, `*`, or `+` without invoking the operators represented by the same characters, search for `\ ?`, `\ *`, or `\ +`.

Two other operators related to word boundaries are `^` and `$`. These correspond to the beginning and the end of paragraphs. They are of greater use to the programmer than the translator and are mentioned here only so the reader knows to search for `\ $` when looking for dollar amounts. Note that this `$` is distinct from the one that we have been using in the Replace field and only applies when searching.

You now know enough about regexes to begin experimenting with practical applications. To start with, you should now be able to decode the date conversion in the introduction to this article. This is great for dates formatted **mm/dd/yyyy**. What about the format **Month dd, yyyy**?

First we must identify the month. We can do this using alternation, separating all of the months with vertical bars thus:

```
January | February | ... December
```

**Jonathan Lukens** is a full-time freelance translator based in Holyoke, MA, working from French and Russian into English. He is ATA-certified (Russian>English) and specializes in legal and financial texts as well as software localization. He has a blog on translation at [www.lukenslanguages.com](http://www.lukenslanguages.com).

A simpler option would be any capitalized word: “a capital letter followed by one or more lowercase letters.” In other words, [ A-Z ] [ a-z ] +.

*Puzzle:* the date and the year can be matched using operators and techniques that you now know. Try to work out the rest of the regex for the date and year on your own and replace them with the European format.<sup>4</sup>

### Where to Learn More

Try out the examples in this article and learn the syntax. I have also assembled a collection of other ready-to-use regexes for Word and Dreamweaver, available at <http://tinyurl.com/2kbro9>. If you are still hungry for more, you might consider learning a scripting language like Perl or Python. Computer programming is not as hard as most people think and the regular expression syntax of these languages outperforms anything available in consumer applications.

Once you learn *how* to use regexes, you will soon identify *where* to use them, eliminating

the need to spend hours picking through a haystack in search of needles. With practice, regular expressions can make your work easier, faster, and more consistent. In a world where increased efficiency means increased profits, this is indeed a skill worth having. □

### NOTES:

- 1 Answer: c) arts. The first character class will not match a vowel and the second will not match a consonant.
- 2 Read the regex aloud: “exactly 5 digits, optionally followed by a hyphen and another four digits.” This pattern will match ZIP codes, with or without the plus-four codes.
- 3 Remembering that numbers are assigned to parentheses from left to right, we can see that \$1 will correspond to the last name and \$2 will correspond to the first name. Thus, we can replace the matches from this regular expression with \$2 \$1.
- 4 Find: ([A-Z][a-z]+) ([0-9]{1,2}), ([0-9]{4})  
Replace with: \$2 \$1 \$3

## ■ A Slightly Faster and Safer Internet FREE!

### A small step for your PC, a giant leap for you

by Carey Holzman

#### You may have never heard of a DNS before, but if you

use the Internet, you use DNS all the time. DNS is short for Domain Name Server and I’m going to teach you what it does and why you should care. To understand what DNS is, let’s start with something similar you are probably already familiar with – cellular phones.

If you’re like most people, you own a cellular phone. You probably have your friends and family programmed into your phone so all you have to do is find their name in your phone’s contact list—or if you have fancy phone, just speak their name—and the phone will dial them for you. In my phone, for example, I have Mom in my contact list. When I want to call Mom, I just search for her name in

my phone and my phone knows to dial the number associated with that name. This is just like what a DNS does.

When you type a web address into your Internet browser, such as Microsoft’s Internet Explorer or Mozilla’s Firefox, the DNS looks up the number associated with that name and then “calls” it. That number is called an IP Address, but it’s really not all that different from a phone number.

For example, if you go to [www.google.com](http://www.google.com) you are really going to 208.67.219.230. But that would be tough to remember, wouldn’t it? Can you imagine if you had to remember 4 sets of numbers for each website you wanted to visit? What a nightmare! To make using the

*Reprint of this article granted by the author, Carey Holzman, who has released it to user groups around the world to use freely in their newsletters.*

**Carey Holzman** is the co-host of the nation's longest running, nationally syndicated radio talk show on computers, Computer America. You can hear Computer America on radio stations across the country or online at [www.ComputerAmerica.com](http://www.ComputerAmerica.com). Carey is also the author of *The Healthy PC*, published by McGraw-Hill. With more than 20 years' experience as a professional PC repair technician, Carey has worked as an instructor at his local college as well as written numerous articles for websites and magazines to help people better understand and take back control of their computers. Carey's personal website is [www.careyholzman.com](http://www.careyholzman.com) and he can be reached directly at [Carey@CareyHolzman.com](mailto:Carey@CareyHolzman.com).

Internet simpler, all you need to know is the site's "name" and your Internet Service Provider's DNS will look it up for you.

The problem is that some Internet Service Providers' DNS computers are out-of-date or over-worked. Others may be well-maintained, but none of them do any filtering for you. If you were to receive an email from your bank asking you to verify your account or encouraging you to log in to your account to resolve some issue, you may actually be

taken to site that looks like your bank in every way, but isn't a bank at all. Instead, thieves have found a way to copy entire web pages, word for word and graphic for graphic, and place it on their own computers. They have a way to make the link you click on in the e-mail take you to their site, instead of to your real bank. When you enter your account number and password, you are effectively telling them your account number and password. If you enter any other information about yourself, such as your first and last name, address, secret PIN for your debit card, social security number, mother's maiden name or other personally identifiable information about yourself, then you can kiss your identity goodbye. You've just given them all the information they need to fill out credit card applications, log into your bank and withdraw all of your funds, or even get a job using your social security number, but never pay any taxes. All of which will have unfortunate and long term repercussions for you that can also be quite expensive and frustrating to fix.

We refer to this process of criminals obtaining your information in this manner as "phishing" (pronounced "fishing"). Perhaps you're under the impression that if you don't click on any links in e-mail (advice often offered as a security measure to prevent

becoming a phishing victim) you'll be safe. But that's not guaranteed. Criminals have figured out ways to manipulate some DNS computers to take you to their site! Imagine if someone could change the number in your cellular phone for Mom, so that each time you went to call Mom, your phone dialed a 1-900 number that charged you \$2 per minute! When crooks manipulate a DNS computer in this manner, we refer to it as "DNS cache poisoning."

Generally speaking, the Internet Service Providers are doing a great job of preventing their DNS computers from being compromised, but how can you be sure? Your Internet Service Provider has a lot of things to do besides maintain their DNS computer. They have e-mail and spam to contend with and they may even provide cable TV, telephone or other services. So how much time and attention do their DNS computers receive?

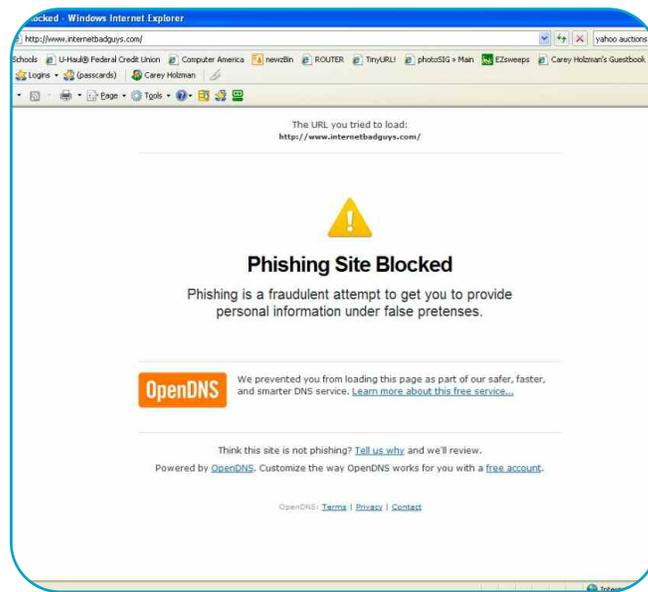
## OpenDNS

One company, [OpenDNS.com](http://OpenDNS.com) does nothing but maintain their DNS computers. In fact, they go a step further by filtering dangerous websites. If you were using their DNS computers, instead of the ones you are using right now, and tried to visit a fraudulent website masquerading as a financial institution in order to collect your personal information, it wouldn't let you go there!

What's more, [OpenDNS.com](http://OpenDNS.com) offers their DNS computers free, for anyone to use, forever. Because their sole focus is on DNS, their computers are extremely fast and are updated constantly with new, dangerous websites to block. You don't have to install any software, sign any license agreements or even tell them who you are to take advantage of their free service.

If you've ever thought your Internet was down, it may have only been that your Internet Service Provider's DNS computers were unavailable for some reason. When the DNS computers are unavailable, your computer is unable to resolve the web addresses and it will appear as though your Internet connection is not working.

[OpenDNS.com](http://OpenDNS.com) has high performance DNS computers in numerous locations around the



On the left, what a fraudulent website might look like without OpenDNS. The image to right is what a fraudulent website will look like with OpenDNS.

world to ensure reliability and speed. They are so proud of how reliable their system is, you can check the status of their entire system at any time. Curious? Go to [system.opendns.com/](http://system.opendns.com/) to see the current status of all of the OpenDNS computers (green = good, red or yellow = bad).

If you choose to register with OpenDNS.com, which is also free, a few more features also become available to you. This may actually appeal more to a business owner than a home user, but if you're curious, it might be worth checking out as well.

OpenDNS.com works with all computers, running all versions of Windows, Linux and even Apple PCs and gaming machines like the Xbox or PlayStation. It can also work within

your router, and if your phone has Internet access, it can work with it as well! Best of all, they provide a free click-by-click walk through (at [www.opendns.com/start](http://www.opendns.com/start)) of how to configure it on all of these devices.

## Start Using OpenDNS

If your computer runs any version of Microsoft Windows, you can quickly configure your computer to use OpenDNS by following a few simple steps that everyone can easily understand at [www.opendns.com/start/windows.php](http://www.opendns.com/start/windows.php). To learn more about how OpenDNS works, visit their FAQ at [www.openDNS.com/faq](http://www.openDNS.com/faq). If you want to learn more details of exactly how DNS works, visit [www.howstuffworks.com/dns.htm](http://www.howstuffworks.com/dns.htm). □

## Language Tech News Submission Guidelines

Submissions should be no longer than 2000 words. (If you have a piece that is longer, contact me and we can discuss options.) Send all articles and items of interest to [lisa@intralingo.com](mailto:lisa@intralingo.com) in Word (.doc or .rtf) format and be sure to include a short bio at the end, ideally containing your e-mail or website contact.

We also gratefully accept reprints of articles you may have published elsewhere, so be sure to offer those as well.

2008 Issues	Submission Deadline	Planned Publication Date
Vol. 2/No. 2	Monday, May 19	June 2008
Vol. 2/No. 3	Monday, August 18	September 2008
Vol. 2/No. 4	Monday, November 17	December 2008

# Trados Tip

by Tuomas Kostiainen

## Substitutions (File > Setup... > Substitutions)

is a very useful but underutilized function in Translator's Workbench. By default, Translator's Workbench recognizes numbers, dates, times, and some measurements and treats them as placeables (non-translatable elements) because these elements need to be included in the translation exactly as they are (such as numbers) or in a localized form (such as dates). Inserting placeables into the target field during translation is easy to do by using the **Get Placeable** buttons on the Translator's Workbench toolbar (or the respective keyboard shortcuts). Note that the substitutions are selected in the Substitutions dialog box (**File > Setup... > Substitutions**) but the substitution localization settings are selected in the **Substitution Localisation** dialog box (**Options > Translation Memory Options > Substitution Localisation**).

In addition to the default substitution of numbers, dates, time, and measurements, Trados also offers two non-default substitution possibilities, namely acronyms and variables. Here acronyms mean anything that's written in uppercase unless the whole sentence is in uppercase. The **Acronyms Substitution** setting (**File > Setup... > Substitutions > Acronyms**) becomes handy when you are working, for example, on instructions where software interface commands or names of buttons and switches are written all in capital letters and don't need to be translated. When "Acronyms" is selected on the "Substitutions" tab, Trados treats all words written in uppercase as placeables and substitutes them automatically, which means that, for example, "Select SEGMENTATION RULES tab" would be a 100% match after "Select ACCESS RIGHTS tab" has been translated.

The other non-default substitution option is called "Variables" (**File > Setup... > Substitutions > Variables**). With this you can define words or phrases, such as proper names, that don't need to be translated by adding them to the "Variable list" box. All these words will then be treated as placeables during translation, and it is very easy to add them to the text using the "Add placeable" commands without ever having to retype them. When you select the **Variables** box, the **Variables** list on the right side becomes accessible. Now you can type the variable names in the list. Note that each

variable has to be on its own line. To enter the next variable, move to the next line by pressing Ctrl+Enter. Variables must appear exactly as in your documents and cannot include any punctuation or hyphens.

Remember to click **OK** after you have finished changing the **Substitutions** settings. Note also that these settings are translation memory-specific and will be active for that translation memory until you change them. Substitution settings, just like all the other setup settings, do not get copied within a TXT export file to a target memory.

The Translator's Workbench User Guide advises you to reorganize (**File > Reorganise**) the translation memory after adding or modifying the variable list if the current memory already contains translation units so that Workbench can process the variables correctly. However, even though I have not always followed this advice, I haven't experienced any problems with my Variable lists.

Substitution localization settings are set at **Options > Translation Memory Options > Substitution Localisation**. Here you can select the variables that you want to have automatically localized and can verify that the suggested formats are correct. You can select the date format separately for short and long dates from the corresponding pull-down menu. The localized date is shown on the right side. Translator's Workbench automatically sets all these localization settings based on the target language of the open translation memory, but if none of the formats offered is what you need, you can edit the choices in the pull-down menus as needed. The same also applies to the time and numbers style. For measurement, you have two choices: metric and U.S. The measurement conversion tool is very handy but far from perfect. For example, it converts lengths expressed in feet and inches (5 ft 6 in.) so that you get two separate cm values; one for the feet (152.40 cm) and another one for the inches (15.24 cm). You need to do the math yourself (152.40 cm + 15.24 cm = 167.64 cm) and round the value appropriately for the use in question (167 cm). □

---

**Tuomas Kostiainen** ([tuomas@jps.net](mailto:tuomas@jps.net)) is an English to Finnish translator and Trados trainer, and has given several Trados workshops and presentations. For more Trados help information, see [www.finntranslations.com/tradoshelp](http://www.finntranslations.com/tradoshelp).