

CHALLENGES IN BI-DIRECTIONAL TEXTS

A guide for translation and DTP professionals
regarding left-to-right and right-to-left languages

Issues, Problems and Solutions

Version 1.0

Tzviya Levin Rifkind

www.sciencetrans.com

tzviya@gmail.com

Introduction

Over my years of work in translation, I have come to realize that people that are unfamiliar with any RTL language have considerable difficulties in understanding the alignment/direction and behavior of these languages, whether monolingual or having both RTL and LTR text in the same sentence. This guide provides some explanations and addresses a variety of issues, thus allowing a person to understand and correct them.

Since I translate from English into Hebrew, I will use the names of these languages as representing a certain direction: English represents a left-to-right (LTR) language while Hebrew represents a right-to-left (RTL) language. In most cases, the explanations apply to other languages with the same direction.

Aside from general explanations, I will provide some techniques and tips for dealing with specific bi-directional issues in Word, Excel and CAT tools. The guide includes explanations of simple and complex directional elements (both text and images), general and specific problems with mathematical symbols and behavioral differences between the three language tools mentioned above. The proposed solutions should allow a person to easily troubleshoot problems in any texts corrupted by bidirectional issues.

Color codes

In this guide, **red** text represents an example while **blue bold** text represents a word/phrase that has a certain definition mentioned in the guide.

Copyright © 2018 by Tzviya Levin Rifkind

All rights reserved. The document may only be distributed in its entirety as a digital copy for personal, non-commercial use.

No part of this publication may be reproduced, distributed or transmitted in any form or by any means, including photocopying, recording or other electronic or mechanical methods, for commercial purposes.

Word¹

As mentioned, Hebrew is a RTL language. Therefore, it should be aligned to the right:

חורשת האקליפטוס, הגשר, הסירה, וריח המלוח על המים.

It is incorrect to align it to the left:

חורשת האקליפטוס, הגשר, הסירה, וריח המלוח על המים.

Punctuation

Hebrew punctuation rules are basically identical to English. A punctuation mark lies immediately after the word with no space between the word and the mark. Yet, due to the direction, "after" is a relative word in this context.

For LTR languages:

Before the text = to the **left** of the text **After** the text = to the **right** of the text

For RTL languages it is the exact opposite:

Before the text = to the **right** of the text **After** the text = to the **left** of the text

Please remember the above definitions!

The punctuation marks in the example we just showed (a comma and a period) lie to the left of (**after**) the word, with no space between a punctuation mark and a word. In this example, even when incorrectly aligned to the left, the visual order² of the sentence is still correct.

Things get complex when a sentence has both Hebrew and English text.

In this guide, any part of non-Hebrew text embedded in a Hebrew sentence is called an **element**. An **element** consists of one unit (a word or a numeric expression), with or without punctuation.

When an **element** has two or more consecutive parts, I refer to it as a **complex element** (e.g., text + punctuation mark). A **complex element** should lie exactly as in the source language.

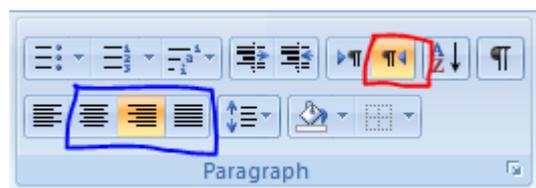
¹ Unless otherwise specified, Office commands are for the 2007 version. Commands may differ slightly for other Office versions.

² In this guide, the term "visual order" describes the way the text looks.

This is the place to say some words about:

Alignment

To avoid issues or if suspecting an issue, first align the text properly. Always make sure that a Hebrew text's direction is RTL and the text is also aligned to the right. The image below shows the paragraph commands from Word (it is the same in other Office programs) where RTL Text Direction (right Ctrl+Shift) is marked with a red frame and the alignment (either *Align Right* [Ctrl+R], *Justify* [Ctrl+J] or *Center* [Ctrl+E]) is marked with a blue frame.



Note: In memoQ, the target text is automatically aligned according to the target language defined in the project.

Tip: Be aware that the left/right* Ctrl+Shift keyboard shortcut changes both direction and alignment. The same behavior applies when the text has LTR direction + left alignment or RTL direction and right alignment and we click only the RTL or LTR direction button, accordingly.

* Left Ctrl+Shift and right Ctrl+Shift behave differently. The left one aligns the text to the left and the right one aligns it to the left. This works only when a Office's Hebrew language package is installed, otherwise, these shortcuts do nothing.

In a bi-directional text, if the direction is correct, the visual order is maintained. Watch out for incorrect alignment.

Tip: An all-English text (e.g., address or citations) standing alone in a Hebrew page (not within a Hebrew paragraph) should always be aligned to the left.

A sentence with one English **element**:

Microsoft has offices in Israel.

In this example, for purposes of demonstration, we will keep the company name in English:

לחברת Microsoft יש משרדים בישראל.

Here is the same sentence with LTR direction and aligned to the left:

לחברת Microsoft. יש משרדים בישראל.

The highlights clearly show that the visual order is different than the logical order³. The two Hebrew parts around the English **element** have switched since Word uses the logical order of the defined direction and alignment (the second sentence has LTR direction and is aligned to the left. So now the first part [yellow] is first from the left).

Tip: To prevent issues, always type the text in a logical order. In other words, never reverse any order to achieve a visual order.

A sentence with two English **elements**:

Microsoft has offices in Israel with Google also having offices in Israel.

לחברת Microsoft יש משרדים בישראל וגם לחברת Google יש משרדים בישראל.

Here is the same sentence with LTR direction and aligned to the left:

לחברת Microsoft יש משרדים בישראל וגם לחברת Google. יש משרדים בישראל.

The visual order of the three Hebrew parts becomes LTR.

However, in this example, as the color highlights show, Word doesn't change the visual order of the period and spaces. Thus, Word maintains the logical order of the language of the last character before the punctuation mark or space.

When such complex text is aligned to the left, the logical order is reversed. Spaces and punctuation marks receive the logical order of the last character. The result is an incorrect visual order, which means that actually the visual order is corrupted⁴.

³ Logical order is the order in which the text is typed (and stored in the computing commands of the file).

⁴ Corrupted text = text in which one or more of the following components is moved in any way from the place and direction it is supposed to be: words, letters, spaces and punctuation marks. A corrupted text is illegible.

The logic here is that the order of both Hebrew and English words is fixed according to the direction in which the row starts (here it is LTR) but the punctuation marks lie according to the direction of the last character (RTL) before them. For this reason, the text is corrupted.

To make life easy for those of you who don't know any RTL language, let us analyze the example:

Punctuation: the period at the end of the sentence lies to the left of the last word (**after** the Hebrew word [green highlighted text]).

Spaces: in the left-aligned example, there is visually no space between **לחברת** and **Microsoft**. However, a sharp-eyed observer can see that the sentence has a small indent to the left. This is the "missing" space (circled below). Using *Show/Hide Mark*, we can see that space lies to the left of **לחברת**. It is **after** the Hebrew word.

⓪ לחברת Microsoft ישראליים וגם לחברת Google ישראליים בישראל

Also, the space we see after **Microsoft** is not really (logically) after **Microsoft** (English), it actually lies **after** **ישראליים וגם לחברת** (Hebrew). We can see it only on a soft copy, when moving the blinking cursor character by character.

A sentence with a **complex element**:

The shares of **Alexion Pharmaceuticals, Inc.** went up considerably.

"Alexion Pharmaceuticals, Inc." is a **complex element**. Therefore, its punctuation marks (which lie **after** the last character of the word) should stay in English (as the **element** itself).

המניות של **Alexion Pharmaceuticals, Inc.** עלו במידה ניכרת.

Tip: When a complex English **element** breaks across rows

המניות של **Alexion**
עלו **Pharmaceuticals, Inc.**
במידה ניכרת.

it makes hard to read. It is advisable to keep it all together

המניות של
Alexion Pharmaceuticals, Inc.
עלו במידה ניכרת.

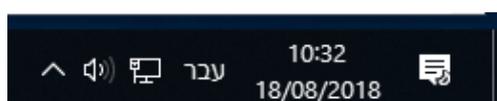
using a non-breaking space (Ctrl+Shift+Space). In case of a hyphen, use a non-breaking hyphen (Ctrl+Shift+-).

Spaces

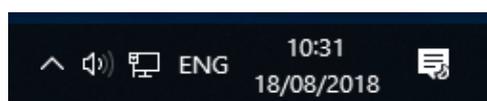
A space (actually "white space") is a blank area between characters and is used as a word divider. A space has a direction but it is impossible to visually determine whether a space is in English or Hebrew. The *Show/Hide Mark* can only show the number of spaces, not their direction.

In order to see the direction of a space, place the blinking cursor at the space and see which direction it points to. You can then also see which language (and hence the direction) shows on Windows taskbar (see below).

Hebrew language in taskbar



English language in taskbar



It is impossible to determine the space direction on a PDF or printed page.

Sometimes, although the text is correctly aligned to the right, there are still spacing issues, usually in a file exported from a CAT tool and not if directly typed into Word:

I love Coldplay's music and also Enya's.

אני אוהבת את המוסיקה של Coldplay וגם של Enya.

The *Show/Hide Mark* can show the spaces themselves but not their direction:

אני אוהבת את המוסיקה של Coldplay וגם של Enya.

The solution for incorrect direction of space is simply replacing the space with a space of the opposite direction.

Tip: A global solution of this issue is easy. Replace each space in the document (or any part you choose, like a sentence or a paragraph) with a space (Ctrl+H > type a space in both blank fields (it doesn't matter in which direction the spaces are typed) > Replace All). Word would insert spaces logically, so the visual is appropriately corrected:

אני אוהבת את המוסיקה של Coldplay וגם של Enya.

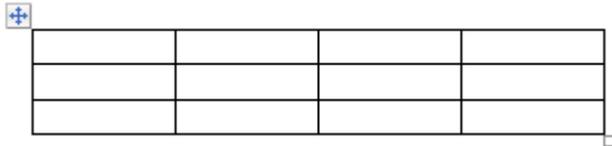
and with the *Show/Hide Mark* activated:

אני אוהבת את המוסיקה של Coldplay וגם של Enya.

Tables

Tables have an additional component direction, setting the general direction of the table itself (table direction).

The blue cross, visible when hovering above the table or making a left click anywhere in a table, marks the top corner of the first column of the table. It can be either on the left (LTR table) or the right (RTL table) side. Below is an LTR table:



When translating from English to Hebrew, there is no need to type the text from the left column (first in English) into the right column (first in Hebrew). Just make sure to reverse the table direction. This will make the table start at the right margin. In addition, the **first** column will be the right one and the **last** column will be the left one. All we have to do is to replace the English text with the Hebrew one and vice versa.

To clarify, **first** is **before** while **last** is **after**.

When working on a Word file in a CAT tool, tables should be automatically reversed during export. When translating directly in Word or if the table didn't flip, it should be done manually (right click in any place in the table > Table properties > check Right-to-Left or Left-to-Right, as desired, under Table direction > OK).

All previous remarks regarding Word also apply to a text within a table cell in Word. Hence, incorrect table alignment is sometimes the only problem requiring fixing in order to correct a visually corrupted table.

Embedded graphics

As mentioned for tables, in Hebrew the layout of columns (e.g., in a leaflet) is RTL. The format in Hebrew is a mirror image of the English one.

When translating from English into Hebrew, the graphics should be adjusted. To do this, we must first identify the location of **before** and **after**.

This is an example text: _____

_____ זה טקסט לדוגמה:

In both sentences above, the line lies **after** the text.



Flowers



פרחים

In both examples above, the picture is placed **before** the text.

For some images, it is not enough to switch locations:



Attention



זהירות

The image here points to something. So, we also need to rotate the image itself (Format picture > 3D rotation > 180° rotation of X axis).



זהירות

Some images



should never be rotated:



Directional bullets

Bullet arrows should be treated similarly. The bullet should lie **before** the text and the arrow should point towards the text.

→ This is an example text.

We can switch the alignment of the text to RTL and type in the Hebrew text instead of the English one. However, the bullet arrow would stay visually the same. Also, CAT tools do not import the bullets but instead keep them as in the exported file as in the source file, without changing their direction:

→ זה טקסט לדוגמה.

To correct this, we should choose a new RTL bullet from the bullet list:

← זה טקסט לדוגמה.

⁵ The different background and sharpness of the flipped image is caused by Word and not subject to choice.

Numbers

Mathematical equations: same as in English (read LTR)

However, sometimes we need to translate the mathematical signs in order to prevent any ambiguity⁶:

...thus, this study arm requires recruitment of 121-75/3=96 new patients...

If we keep the mathematical formula intact:

...לכן, צריך לגייס לזרוע זו של המחקר 121-75/3=96 מטופלים חדשים...

The words "new patients" relates to "96" but, in the above Hebrew translation, the term and its number are too distant from each other.

Therefore, it is better to translate the mathematical functions into words:

...לכן, צריך לגייס לזרוע זו של המחקר 121 פחות 75 חלקי 3 שהם 96 מטופלים חדשים...

Percent signs: As with equations, a number with a % sign is written exactly as in English, i.e., LTR. First the number followed by the % sign to the right of it, with no space in-between. It is opposite of the direction of the Hebrew language although visually it may seem that the % sign comes before the number (to the right of the number). To the best of my knowledge, the Arabic used in Israel (when using western numbers) places the number first and then the % sign (i.e. to the left of the number: %97) and reads RTL.

Range of numbers: The starting and ending number should be written RTL as determined in 1940 by Va'ad Halashon, the body that preceded the Academy of the Hebrew Language. However, the Academy of the Hebrew Language's decisions are not mandatory. As a result, many Hebrew speakers have always written it LTR, as in English. However, in recent years, people have increasingly applied the Academy rule, writing ranges in the correct direction, RTL. Yet, many people continue to write it LTR, creating a lack of uniformity. The best we can do is to maintain consistency in a file or a project.

Hyphen, dashes and minus sign: There is another influencing factor related to the behavior of Word and CAT tools. This factor is a group of small, similar and confusing lines, specifically the hyphen, dash and minus sign.

The two that are of interest to us here are the hyphen (-), which is narrow, and the en dash (–), which is wider. There is no dedicated key on a keyboard for the en dash. So, using it requires us to remember the appropriate keyboard shortcut: Alt+0150 on the numeric keypad. A newer keyboard shortcut is left⁷ Ctrl+- (there is a new standard for the Hebrew keyboard by The Standards Institution of Israel). Historically, due to the lack of a dedicated keyboard key for the en dash, many people replace it with a hyphen, which does have a key and is easier to type.

⁶ Indeed, there is a better way to phrase this example in English. However, the example is based on text I encountered and helps to demonstrate the issue.

⁷ Left Ctrl results in a different character than right Ctrl.

The hyphen is a feature that connects two parts while the en dash is a separating feature. However, the correct Hebrew hyphen lies at the top of the row (־), not in the middle (-) as in English. The new keyboard shortcut for it is right⁸ Ctrl+-. It is usually used in books, not in ordinary documents.

So basically, for reasons of practicality, Hebrew speakers got used to using a hyphen (־) to serve as both the upper hyphen (־) and the en dash (–).

In Word, when using a hyphen (־) in a range of numbers, the range automatically flows LTR⁹:

I will arrive in 5-6 days.

אני אגיע בעוד 5-6 ימים.

So, in order to have a RTL range, we should use the en dash:

אני אגיע בעוד 6–5 ימים.

There are two other ways to produce the correct visual order of numbers but they are incorrect:

- Using a hyphen and typing the numbers in a reverse order:

אני אגיע בעוד 6-5 ימים.

Please remember the **Tip** from page 4. Reverse typing does not support the logical order and might cause issues if the text is further processed in other formats.

- Inserting a space before and after the hyphen:

אני אגיע בעוד 5 - 6 ימים.

As mentioned above, it is incorrect to insert these spaces.

⁸ See previous footnote.

⁹ This also happens in Trados (2007) when a hyphen is properly used to connect a Hebrew character and a non-Hebrew character. For example, "Friday the 13th" is "יום שישי ה-13" ("ה" is "the"). In Trados, the visual of "13-ה" would be "ה-13" (the number after the hyphen will lie to its right). Linguists that are not familiar with the use of directional marks (see page 17) solve this issue by using a space ("ה-13"), which is incorrect in Hebrew.

Visual scales

A scale containing descriptive text for each value is written from right to left, the same direction in which Hebrew is written:

יעיל ביותר = 5 (extremely effective) יעיל מאוד = 4 (very effective) יעיל = 3 (effective) יעיל במקצת = 2 (somewhat effective) כלל לא יעיל = 1 (not at all effective)

A scale containing only numbers for each value is somewhat confusing.

The preferred way is a RTL layout, the same as in Hebrew:

יעיל ביותר (extremely effective) 5 4 3 2 1 כלל לא יעיל (not at all effective)

However, since mathematics is written LTR and, as mentioned in the previous page, some people write number ranges LTR, an LTR layout of scales is also valid:

כלל לא יעיל (not at all effective) 1 2 3 4 5 יעיל ביותר (extremely effective)

So, from a global point of view, consistency must in any case be maintained within a file or a project or for a certain client.

The stubborn inch mark

The inch expression ("") should be treated like a mathematical expression (see the "Numbers" section above) and remain as in English.

When typing a Hebrew sentence, we naturally type the numbers RTL (the language shown in Windows taskbar is HE [RTL] with the numbers being automatically written LTR). Unfortunately, it does not work with the inch expression, which in Word should be typed in English.

In the following example¹⁰

A 500ml bag with an intra-duodenal feeding tube 10 Fr (0.13") X 120cm (47")

the content within the parentheses was typed RTL:

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסריונית בקוטר 10 Fr (0.13") ובאורך 120 ס"מ (47")

causing corruption of the content within parentheses.

However, when the content within the parentheses was typed LTR, the inch mark fully cooperates, with no corruption of the text:

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסריונית בקוטר 10 Fr (0.13") ובאורך 120 ס"מ (47")

Inch marks are especially stubborn. The above solution does not work in memoQ. Even if typed in English, the content within the parentheses is corrupted in the exported file with the inch mark remaining to the left of the number. It should be corrected using a directional mark.

Directional marks and how to use them in memoQ are discussed in the CAT tools chapter (page 16).

¹⁰ This example includes a unit of measure (Fr). An explanation follows in the next chapter.

Units of measure

Units of measure (e.g., L/min, cmH₂O and MHz) should be translated into Hebrew. There is no linguistic justification for leaving them untranslated as some customers request. Furthermore, translation of units of measure prevents format and DTP issues.

Let's refer to a numeric expression + unit of measure (e.g., 60 MHz) as a "block".

In that block, the unit of measure is always placed **after** the numeric expression (e.g., 100; 50,000 and 0.76 ± 0.1) with the number and the unit of measure separated by a space.

Usually the Hebrew reader thinks in Hebrew and expects to find the unit of measure after the numeric expression, as the Hebrew sentence structure dictates (without "jumping" back and forth between RTL and LTR directions).

In any case, we handle the block according to the Hebrew language rules, meaning the unit of measure is placed to the left of the numeric expression even if the unit of measure is left in English.

By contrast, many think that a block should be handled according to the rules of the original language (English) as shown in the 2nd alternative version of the examples below. This time the **after** is to the right of the numeric expression since the English language rules apply. The advocates of applying source language rules to that block argue that the eye is trained to see it as one unit. In this era of global communication, it is common to find untranslated text of one language (e.g., English) in another language (e.g., Hebrew).

The Academy of the Hebrew Language hasn't discussed this subject and thus currently has no official position.

The examples below mostly represent the approach I tend to take but it is not binding. As mentioned for range of numbers (page 10 above), the guiding principle is to maintain consistency in a file or a project.

Example 1: 120 kHz to 60 MHz

Preferred: 120 קילוהרץ עד 60 מגהרץ

Alternative 1: 120 kHz עד 60 MHz

(In this example, there is no need for any directional marks (explanation will follow shortly) since memoQ/Word uses the logical order, which matches the visual order).

Alternative 2: 60 MHz עד 120 kHz

Example 2: ± 2 kV for power supply lines

In this example, the numeric expression is actually a mathematical expression (± 2). As mentioned above, mathematical expressions are written as in English, LTR.

Preferred: ± 2 קילוולט לקווי אספקת חשמל

Alternative 1: ± 2 kV לקווי אספקת חשמל

(please see detailed explanation on page 20)

Alternative 2: ± 2 kV לקווי אספקת חשמל

Example 3: Maximum limited pressure: 120 ± 5 cmH₂O

Here the numeric expression is mathematically more complex (120 ± 5):

Preferred: לחץ מוגבל מרבי: 120 ± 5 ס"מ מים

Alternative 1: לחץ מוגבל מרבי: 120 ± 5 cmH₂O

Alternative 2: לחץ מוגבל מרבי: 120 ± 5 cmH₂O

CAT tools

The compatibility level for RTL languages differs among CAT tools. Sometimes the incompatibility is already visible in the CAT tool file while other times it appears only in the clean file after export.

Incompatibility includes incorrect reversing of **elements** such as the following:

- Parentheses
- \leq and \geq signs
- Direction of parts of **complex elements** (e.g., "Inc.").
- % sign in Hebrew (see image from memoQ below. Segments 17 & 19 are without any correction; segments 18 & 20 are corrected. I will shortly explain how to correct),

17.	97%		%97	0%	✓
18.	97%		97%	0%	✓
19.	24 (35%) of them were males.		.מתוכם היו בנים. 24 (%35)	0%	✓
20.	24 (35%) of them were males.		.מתוכם היו בנים. 24 (35%)	65%	✓

The use of the % sign varies among RTL languages. In Arabic (so I was told), it is written "%97". First we read the number (from the right) and then the % sign with no space in between (actually, as shown in row 17 target in the above memoQ image). In Hebrew, it is handled the same as mathematical elements. It is written 97% and reads LTR as in English (as in row 18 target in the above memoQ image).

To clarify, in English, Hebrew and Arabic first we read the number and then say "percent" (%). In English and Arabic it is written in the respective language direction. Yet, in Hebrew, it is written in the opposite direction, LTR.

Reversing occurs in memoQ or in the exported file. In Word, if we want to type a sentence like:

לחברת Microsoft יש משרדים בישראל.

we type לחברת (RTL) and revert the blinking Word cursor and type Microsoft (LTR). Then we revert the blinking cursor again and type יש משרדים בישראל. (RTL). We can revert the blinking cursor again but it is easier just to press the *End* key. This will put the blinking cursor in the logical direction, which matches the direction and alignment of the first character of this sentence, i.e., RTL (defines the language).

You can revert the blinking Word cursor (i.e., changing the language) in one of two ways: (i) changing the language in the Windows taskbar (see page 6) according to your Windows version and (ii) using Alt+Shift.

Please note that pressing the *End* key will not revert the blinking cursor in memoQ. Also, punctuation marks in a CAT tool target segment receive the logical direction of memoQ's defined target language. A sentence as in the example we used earlier:

המניות של Alexion Pharmaceuticals, Inc. עלו במידה ניכרת.

would never have a correct visual order without additional aid. This aid is called **Directional Marks**.

Directional Marks

Directional marks are zero-width marks (Unicode control characters). They influence the bi-directional ordering exactly as a character would influence it. However, they are invisible. There are two types of directional marks:

RLM (right-to-left mark): Sets the logical direction from right to left.

LRM (left-to-right mark): Sets the logical direction from left to right.

To correct the direction of the first character, add a directional mark just **before** this character.

To correct the direction of the last character, add a directional mark just **after** this character.

Characters (letters, numbers, parentheses, punctuation marks and spaces) receive the logical order of the character just before them. Therefore, if one character is in one language while the next character is in a language of the opposite direction, the visual order should be corrected by inserting the appropriate directional mark.

When translating in a CAT tool without using the appropriate directional marks, the exported result might be visually corrupted, with the corruption sometimes already visible in the CAT tool.

[For a detailed explanation on viewing the invisible directional marks, see page 25.](#)

Visual corruption in exported file – The logical order is correct, but the incorrect direction of the spaces, parentheses and inch marks results in a corrupted visual*:

A 500ml bag with an intra-duodenal feeding tube 10 Fr (0.13") X 120cm (47")

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסריונית בקוטר 10 Fr (0.13") ובאורך 120 ס"מ (47")

Here is the same example with visible non-printing characters*:

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסריונית בקוטר 10 Fr (0.13") ובאורך 120 ס"מ (47")

The same example with spaces, parentheses and inch marks direction corrected (all are now in the Hebrew direction, RTL)*:

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסריונית בקוטר 10 Fr (0.13") ובאורך 120 ס"מ (47")

The same corrected example with visible directional marks*:

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסריונית בקוטר 10Fr (0.13") ובאורך 120 ס"מ (47")

The same corrected example with visible directional and paragraph marks*:

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסריונית בקוטר 10Fr (0.13") ובאורך 120 ס"מ (47")

* The reason for the lack of color highlights in the examples with visible directional marks is a Word 2016 bug (see page 26). As an aid to overcome this inconsistency, see below a capture from memoQ showing the same example with visible directional and paragraph marks and highlights:

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסריונית בקוטר 10Fr (0.13") ובאורך 120 ס"מ (47")

If an item (e.g., space, parentheses, comma, period and colon) is reversed, (English [LTR] instead of Hebrew [RTL] or vice versa), use Find and Replace (Ctrl+H) > Type the item in both **Find What** and **Replace With** fields (it doesn't matter if typed LTR or RTL!) > Replace All. The new items will be automatically inserted in the correct direction.

Tip: If such correction is needed for more than one type of item in a file, replace the English style with a Hebrew style as follows:

1. Find and Replace (Ctrl+H)
2. More
3. Place the cursor at the **Find What** field, open the Format dropdown menu, from Language choose your source language (e.g., English). Don't type anything.
4. Repeat for **Replace With** field and select your target language (e.g., Hebrew). Again, don't type anything.
5. Replace All.

The following types of **elements** require directional marks:

- Percent sign in Hebrew with no preceding RTL text in that row [97%] (if this **element** has a preceding RTL text [כמעט 97% מהם היו שם], the visual order is maintained)
- Numbers separated by a hyphen, such as a phone number [+972-4-5555555]
- Mathematical expressions [120 ± 5]
- Temperatures [4°C]. (see detailed explanation on page 25).
- Names of corporations with abbreviated corporate status marker [Alexion Pharmaceuticals, Inc.; Teva, Ltd.¹¹].
- ®, ©, ™ and any other symbols connected to an LTR word (Coca-Cola®).
- An **element** in a Hebrew text containing an LTR part [e.g., English text, numbers, symbols, etc.] in parentheses followed by another such part [Absolute neutrophil count (ANC) 1.25×10^9 cells/L¹²]. The parentheses retain the logical order of the previous character. If that character is a number/symbol or an English word, the parentheses logical order will be LTR while they should be RTL since the **element** is in Hebrew sentence.

Sometimes the same directional mark is needed both **before** and **after** an **element**.

Do you remember the stubborn inch expression mentioned earlier?

In memoQ, an inch expression requires directional marks both **before** and **after** it. Here is the example with visible directional marks (copied from memoQ):

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תרופות בקוטר 10 (0.13")
ובאורך 120 ס"מ (47")

Examples for use of directional marks:

Example 1:

Absolute neutrophil count (ANC) 1.25×10^9 cells/L.

ספירת ניטרופילים מוחלטת (ANC) > 1.25×10^9 תאים לליטר.

RLM is used to keep the ">" symbol in the RTL order (left side of the mathematical expression is greater than the right side) followed by an LRM to keep the mathematical expression (1.25×10^9) LTR. It looks like the RLM and LRM are far from each other, but actually they are only one space apart. Also, the "9" and the next Hebrew word (תאים) are also one space apart.

¹¹ I included the English source of the example since in no way I could make Word produce a correct visual for complex Hebrew text in between the English text.

¹² Ibid.

Example 2 (the example from page 15):

±2 kV for power supply lines

±2 kV לקווי אספקת חשמל

Here the memoQ/Word logical order is not enough to match the visual order since "±2" is a **complex element**. Therefore, we need directional marks.

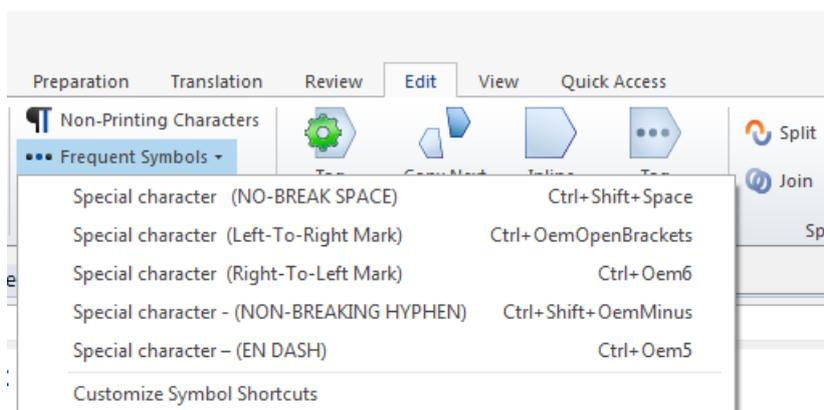
Two directional marks are needed: An LTR mark before "±2" and a RTL mark after "±2". The first keeps the sign "±" **before** the number while the latter maintains the unit of measure **after** that **complex element**. The same applies to the next example in which the **complex element** is more complex $[120 \pm 5]$ ¹³.

A strange behavior that I have noticed is that if we copy a sentence with directional marks like the one above from memoQ, paste it to Word and delete the directional marks, the visual order is maintained:

שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תהסרוגנית בקוטר 10 (0.13") ובאורך 120 ס"מ (47")

The same behavior occurs in XTM editor. I am not sure what the underlying explanation for this phenomenon is.

Directional marks in memoQ are found under *Frequent Symbols* in the Edit tab. It is also possible to set a keyboard shortcut for them at the bottom of the same menu.



¹³ In "±2" there is no space between the sign and the number since the sign describes the value direction of the number. In "120 ± 5", there is a space before and after the sign since this element is a mathematical function.

Excel

False corruption

In the tables below, the shaded rows represent the visually correct sentences.

A corruption of a sentence in a cell might be only a visual illusion. You can notice the visually corrupted elements by comparing the correct sentence to the three example sentences above it (same elements have same colors). The only difference between all four sentences in each table is the text alignment and direction. All are typed exactly the same.

Excel:

	A	B	C	D
1	English	Hebrew	Text direction	Text alignment
2	A 500ml bag with an intra-duodenal feeding tube 10 Fr (0.13") X 120cm (47")	-שקית בנפח 500 מ"ל עם צינורית האכלה תוך-10 ובאורך 120 ס"מ (0.13") Fr תריסריונית בקוטר (47")	LTR	left
3		-שקית בנפח 500 מ"ל עם צינורית האכלה תוך-10 ובאורך 120 ס"מ (0.13") Fr תריסריונית בקוטר (47")	LTR	right
4		-שקית בנפח 500 מ"ל עם צינורית האכלה תוך-10 ובאורך 120 ס"מ (0.13") Fr תריסריונית בקוטר (47")	RTL	left
5		-שקית בנפח 500 מ"ל עם צינורית האכלה תוך-10 ובאורך 120 ס"מ (0.13") Fr תריסריונית בקוטר (47")	RTL	right

Word:

English	Hebrew	Text direction	Text alignment
A 500ml bag with an intra-duodenal feeding tube 10 Fr (0.13") X 120cm (47")	-שקית בנפח 500 מ"ל עם צינורית האכלה תוך-10 ובאורך 120 ס"מ (0.13") Fr תריסריונית בקוטר (47")	LTR	Left
	-שקית בנפח 500 מ"ל עם צינורית האכלה תוך-10 ובאורך 120 ס"מ (0.13") Fr תריסריונית בקוטר (47")	LTR	Right
	-שקית בנפח 500 מ"ל עם צינורית האכלה תוך-10 ובאורך 120 ס"מ (0.13") Fr תריסריונית בקוטר (47")	RTL	Left
	-שקית בנפח 500 מ"ל עם צינורית האכלה תוך-10 ובאורך 120 ס"מ (0.13") Fr תריסריונית בקוטר (47")	RTL	right

Excel might influence the visual order even if we just click in a cell without actually changing anything or only change the column width. In order to know if a cell's visual is correct, we can stretch the cell such that its whole content fits in a single line:

	A	B	C	D
1	English	Hebrew	Text direction	Text alignment
2	A 500ml bag with an intra-duodenal feeding tube 10 Fr (0.13") X 120cm (47")	10 תוך-תריסרונית בקוטר Fr (0.13") שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסרונית בקוטר 10	LTR	left
3		10 תוך-תריסרונית בקוטר Fr (0.13") שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסרונית בקוטר 10	LTR	right
4		(47") שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסרונית בקוטר Fr 10 (0.13") ובאורך 120 ס"מ (47")	RTL	left
5		(47") שקית בנפח 500 מ"ל עם צינורית האכלה תוך-תריסרונית בקוטר Fr 10 (0.13") ובאורך 120 ס"מ (47")	RTL	right

This method is only practical for very short texts. If we want to avoid changing Excel's cell settings (and potentially damaging its content), we can check the visual using the preview of a CAT tool before export or, if the text is already exported to Excel, just copying it to Word. If we copy a single cell, only the text is copied. If we copy multiple cells, the table lines are also copied and create a table in Word. Either way is fine. We can also click in a cell and copy all its content. However, this method is not recommended if we don't know that language or if the cell contains a lot of text because we might lose some of it.

Corruption

A sentence like

המניות של Alexion Pharmaceuticals, Inc. עלו במידה ניכרת.

will be visually corrupted in an Excel cell. Punctuation marks, including the comma and period, that lay **after** an English character maintain the English direction. Only an LTR directional mark will put it back to order. Below is the same sentence in Excel. The top one is without a directional mark with the period visually misplaced (arrow). The second one has a directional mark after the "Inc." with the period in place (arrow).



Note: If we copied a logically and visually correct sentence (as the above example) from Word and pasted it into Excel, the visual would be compromised (the period location would be as in the first example sentence). If we then copied the same sentence back from Excel to Word, the corruption caused in Excel would be retained!

If we transferred that sentence in the opposite direction, namely from Excel into Word, again, the visual would be compromised although the directional mark would still be there. If we then copied the same sentence back from Word to Excel, no change in trend would occur. The corruption caused in the transfer out from Excel would be retained.

Inserting directional marks

In **Excel** as in Word, the *Add directional marks* function can be found under the Insert tab > Symbols [Office 2016] > Symbol > Special Characters > Left-to-right Mark / Right-to-Left Mark. Insertion of one mark requires at least 5 mouse clicks. Unlike in Word, the *Special Characters* window must be closed after each insertion. Otherwise the sheet content becomes inaccessible (accompanied by a sound of dissatisfaction).

Another option is to copy the text to **Notepad**, add the directional marks and then copy the updated text back to Excel. In order to add the directional marks, first make a right click and choose for Hebrew a right to left reading order (when closing Notepad, the text display always return to LTR). Then right click anywhere in the opened file and select the required directional mark (LRM or RLM) from the list under *Insert Unicode Control Character*. From that same place you can also make directional marks visible by choosing *Show Unicode Control Character*.

The keyboard shortcuts (numeric keypad) in **Notepad** and **Word** are Alt+0254 for RLM and Alt+0253 for LRM. Please note these shortcuts are only valid for RTL languages (in other languages, these shortcuts create different signs).

The most efficient way is to copy the text to **memoQ**, add the directional marks (using the appropriate keyboard shortcut is recommended) and then copy the updated text back to Excel (eliminating the need to save, export or take any further action in memoQ). The preview functionality in memoQ (and other CAT tools) is very useful in verifying the correct visual order even before pasting the text back in Excel.

The traitorous nature of temperature expressions

Temperature expressions ($X^{\circ}\text{C}$ and $X^{\circ}\text{F}$) require special attention. They might seem visually correct in memoQ but will not maintain the visual in the clean file or might be logically correct but will not be visually correct in Excel.

Ship at 2°C to 4°C (35°F to 39°F).

Incorrect visual in Excel:



יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

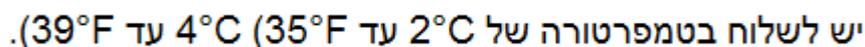
while this is the correct visual:

יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

The temperature expressions themselves are visually correct. Therefore, we might think they are also logically correct with only some space issues. However, when this sentence is copied to Word the problem is obvious:

יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

A similar false visual happens in memoQ:



יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

and in Notepad:

יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

Knowing that traitorous nature of temperature expressions, always make sure to insert an LRM before them.

Here you can see one after the other, the correct visual of this example:

יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

And the same sentence with the appropriate directional marks that we used:

יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

Email

Make sure to align Hebrew sentences to the right, especially those that contain English **elements**. Otherwise, the text will become visually corrupted.

Viewing the invisible directional marks

Sometimes it is helpful to see the directional marks.

In **memoQ**, a tiny red dot (marked below with blue arrows) shows where the mark is placed. However, it does not show whether it is RLM or LRM.



יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

We can better see them, including their direction, using the non-printing characters (same as in Word):

יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

We can see that the mark points to the left or right. Please note the two adjacent marks to the left of "4". The vertical line is thicker with two "heads" on the top pointing in different directions.

This is how the same example looks like in Word after being copied from memoQ:

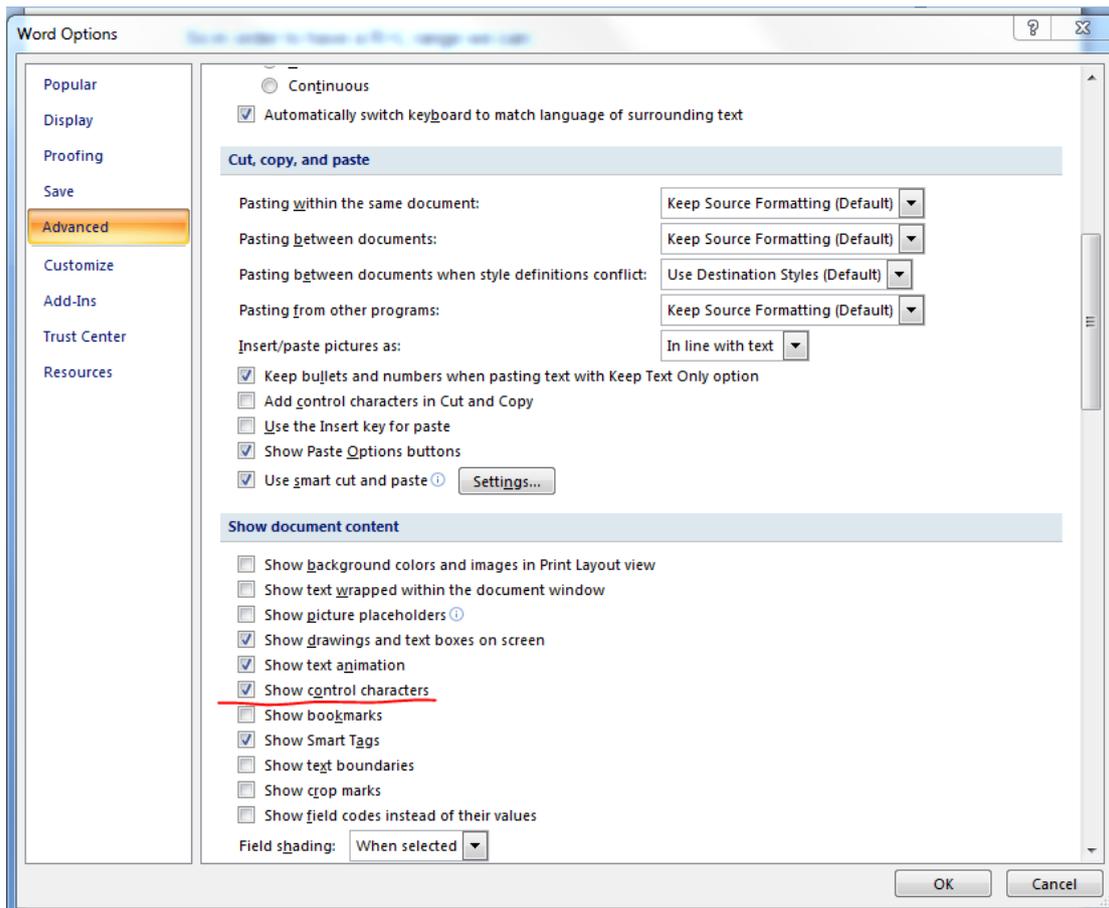
יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

This is how it looks in Word with visible directional marks:

יש לשלוח בטמפרטורה של 2°C עד 4°C (35°F עד 39°F).

The directional marks are the  (RLM) or  (LRM) marks.

In order to show the directional marks in **Word**, check the box for *Show control characters* as seen below:



* The above image is from Word 2007. The *Word Options* window looks similar in Word 2016. During my work on this guide, I switched from Word 2007 to Word 2016. When I checked the *Show control characters* in Word 2016, I was unpleasantly surprised to find out that this feature, although present in the *Word Options*, does not work.

Highlights of characteristics of the Hebrew language

Capitals

The Hebrew language has its own letters. There are 22 letters, of which 5 have a second form used when that letter is the last letter in the word. There are no capital letters as in English. When capitalization in English is used for highlighting (both first capital or all capitals), the equivalent Hebrew highlight options are to put the text between parentheses, either double or single, as long as consistency is maintained throughout the file, or use **Bold**, underlined or both **bold and underlined** text. The choice depends on the pre-existing highlights in the document.

Diacritics

Hebrew doesn't have vowels. The way to indicate the correct pronunciation of a letter is with diacritics (nikkud [<https://en.wikipedia.org/wiki/Niqqud>]). It can be placed below, above or after a letter. Some diacritics include a letter. There is also a special diacritic that is in the middle of the letter. The spelling of the same word might be different between scripts with or without diacritics ("חֶפֶז" vs. "חִפֵּז" [means "looked for"]). In everyday life, we use the Hebrew script without diacritics. Since there are words with the same spelling but different diacritics, we must sometimes use diacritics on that word in order to avoid confusion.

Appendix – list of keyboard shortcuts in this guide

Action	Shortcut	Comment
LTR Text Direction	Left Ctrl+Shift	
RTL Text Direction	Right Ctrl+Shift	
Align Right	Ctrl+R	
Justify	Ctrl+J	
Center	Ctrl+E	
Replace all	Ctrl+H	
En dash (–)	Alt+0150 (Hebrew only); Right Ctrl+-	
Hebrew upper hyphen (-)	Left Ctrl+-	
Revert blinking cursor direction	Alt+Shift	
Find and Replace	Ctrl+H	
RLM directional mark	Alt+0254	Word and Notepad
LRM directional mark	Alt+0253	Word and Notepad
Non-breaking space	Ctrl+Shift+Space	
Non-breaking hyphen	Ctrl+Shift+-	

Acknowledgements

First, I wish to thank my beloved husband, Stephen Rifkind (<http://www.gaguzia-translations.com>), that without his continuous support, help, editing, corrections, love and many cups of tea, I wouldn't have completed this guide.

I also wish to thank the following people for reviewing the guide and commenting: Uri Bruck, Sari Nilgun and Daniel Castillo.

Your feedback is very welcome! Please contact me at tzviya@gmail.com.

Check periodically for the recent version of this guide, which is available at <http://www.sciencetrans.com/Bi-Di-Guide.html>