ZWNJ
Confusing Cases

Arabic Script IDN Working Group
Version 1.0 10 April 2008

Revision History:

Date	Version	Changes	Changed by
10 April 2008	1.0	Created	A. Al-Zoman

This document lists some cases when the ZWNJ control character has no effect on the shape of the character it follows and hence it cause confusion with the same character sequence when ZWNJ is not used.

Usage of ZWNJ
>   "The Unicode Standard provides a user-selectable formatting codes: U+200C zero width non-joiner and U+200D zero width joiner. The use of a non-joiner between two letters prevents those letters from forming a cursive connection with each other when rendered. Examples include the Persian plural suffix, some Persian proper names, and Ottoman Turkish vowels." … The Unicode Standard 5.0 - Chapter 8 Middle Eastern Scripts

Format (sequence of typing … right-to-left)
>   <Succeeding Character> ZWNJ <Preceding Character>
>   SChar = Succeeding Character , PChar = Preceding Character>

1. When the shapes of PChar with or without left joining are similar

This is valid for characters with the TAH shape (ط) which include the following characters:
- TAH U+0637 ط
- ZAH U+0638 ظ
- TAH with three dots above U+069f ڟ

Examples not using ZWNJ	Examples not using ZWNJ
طبل	طبل
input[0] = U+0637 input[1] = U+0628 input[2] = U+0644	input[0] = U+0637 input[1] = U+200c input[2] = U+0628 input[3] = U+0644

2. When SChar is a digit

Since digits are not right joining characters, then having ZWNJ before them has no effect.
- European U+0030..U+0039 0123456789
- Arabic-Indic U+0660..U+0669 ٠١٢٣٤٥٦٧٨٩
- Eastern Arabic-Indic U+06F0..U+06F9 ۹ ۸ ۷ ۶ ۵ ۴ ۳ ۲ ۱ ۰

Examples not using ZWNJ	Examples not using ZWNJ
ب١	ب١
input[0] = U+0628 input[2] = U+0661	input[0] = U+0628 input[1] = U+200c

	input[2] = U+0661

3. When SChar is a dash or a dot

"Dash" and dot (used as label separator) are not right joining characters, having ZWNJ before them has no effect.
- Dash U+002d
- Dot U+002e

Examples not using ZWNJ	Examples not using ZWNJ
ب-	ب-
input[0] = U+0628 input[2] = U+002d	input[0] = U+0628 input[1] = U+200c input[3] = U+002d

Examples not using ZWNJ	Examples not using ZWNJ
ب.	ب.
input[0] = U+0628 input[3] = U+002e	input[0] = U+0628 input[1] = U+200c input[3] = U+002e

4. When SChar is Hamza and High Hamza

Hamza and High Hamza are not right joining characters. Therefore, having ZWNJ before them has no effect.
- Hamza U+0621
- High Hamza U+0674

Examples not using ZWNJ	Examples not using ZWNJ
بء	بء
input[0] = U+0628 input[2] = U+0621	input[0] = U+0628 input[1] = U+200c input[2] = U+0621

Examples not using ZWNJ	Examples not using ZWNJ
بٴ	بٴ
input[0] = U+0628 input[2] = U+0674	input[0] = U+0628 input[1] = U+200c input[2] = U+0674