Comparison of pairs of utterances

<table>
<thead>
<tr>
<th>Filename</th>
<th>UttID</th>
<th>Speaker</th>
<th>Utterance</th>
<th>UttID2</th>
<th>Speaker2</th>
<th>Utterance2</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 ann10614.cha</td>
<td>14</td>
<td>MOT</td>
<td>oh ein bus ein kleiner polizeibus.</td>
<td>15</td>
<td>MOT</td>
<td>noch mehr?</td>
</tr>
<tr>
<td>19 ann10614.cha</td>
<td>19</td>
<td>MOT</td>
<td>ja guck mal.</td>
<td>20</td>
<td>MOT</td>
<td>die kamera die dreht sich wieder mit 'ne?</td>
</tr>
<tr>
<td>20 ann10614.cha</td>
<td>20</td>
<td>MOT</td>
<td>die kamera die dreht sich wieder mit 'n?</td>
<td>21</td>
<td>MOT</td>
<td>is da noch mehr drin Anna?</td>
</tr>
<tr>
<td>27 ann10614.cha</td>
<td>27</td>
<td>MOT</td>
<td>und noch ein bus.</td>
<td>28</td>
<td>MOT</td>
<td>is da noch mehr drin Anna?</td>
</tr>
</tbody>
</table>

Retrieval of repetitive patterns from large collections of child language data
Logical distinction of different overlaps:

I. Completely identical:

II. Completely distinct:
Logical distinction of different overlaps

III. Variations:

- Be careful with that
- Robb
- Across the street, by the yellow house
- There is a ball
- I need you to take care of that
Logical distinction of different overlaps

III. Variations:

- Look, we are getting closer to the house.
- Look, it is over.
- It is over there, close to the house.
- Look, here, close to the house.
- By the river, the ball is close to the house.
- Here, with the red roof, close to the house.
- Look, there, close to the house.
Logical distinction of different overlaps

- Defining exclusive and exhaustive categories
- Any utterance can be assigned to one and only one category

<table>
<thead>
<tr>
<th>Filename</th>
<th>Uttd</th>
<th>Speaker</th>
<th>Utterance</th>
<th>Uttd2</th>
<th>Speaker2</th>
<th>Utterance2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fal20107.cha</td>
<td>180</td>
<td>MOT</td>
<td>komm mal her.</td>
<td>181</td>
<td>MOT</td>
<td>komm.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Identical</th>
<th>Distinct</th>
<th>Inclusion</th>
<th>OL Beginning</th>
<th>OL End</th>
<th>OL Beginning + End</th>
<th>OL Middle</th>
</tr>
</thead>
<tbody>
<tr>
<td>FALSE</td>
<td>FALSE</td>
<td><strong>TRUE</strong></td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
<td>FALSE</td>
</tr>
</tbody>
</table>
Identical:
1. All words of utterance 1 appear in the exact same order in utterance 2 (and vice versa)

```r
n <- 1
for (i in UT1NORM) {
  WordlevelDF$AbsIdentity[n] <- identical(UT1NORM[[n]], UT2NORM[[n]])
  n <- n+1
}
```
Overlap Beginning:
1. First word is the same
2. Last word is not the same
3. No utterance is included in the other

```r
n <- 1
for (i in UT1) {
  WordlevelDF$OLBeg[n] <- isTRUE(is.element(UT1NPWL[[n]][1], UT2NPWL[[n]][1]) &
                                     is.element(UT1NPWL[[n]][length(UT1NPWL[[n]])],
                                     UT2NPWL[[n]][length(UT2NPWL[[n]])])) == FALSE &
  WordlevelDF$Inclusion[n] == FALSE)
  n <- n+1
}
```