Abstract
We attack the problem of aligning words from pairs of bilingual sentences, rather than the well-known, and somewhat easier, problem of aligning sentences. The method that we develop is based on the use of bilingual dictionaries, having supposed that lemmatization has taken place. We first show that this method performs poorly in terms of silence and noise. To improve its performance we introduce syntactic dependency relations between the words in each of the two sentences considered. In this sense the syntagmatic level comes to the rescue of the paradigmatic level at which the alignment actually takes place.

I. Introduction
Given that two sentences F and E are translations of each other. Is there a simple method for aligning the words in each sentence? In other words is word alignment algorithmically simple to implement? We shall see that this problem is extremely delicate, even to be done by hand. To convince oneself, one need merely attempt it in order to see how quickly the choices pass from trivial to complicated. The difficulties stem from there not always being a simple one-to-one correspondence between the words of the two sentences. One word may correspond to many words (an expression); in other cases, one or many words may correspond to no other words. On the other hand, word order is rarely maintained, and to top things off, different syntactic status create complicated pairings when they do exist.

II. Conventions and restrictions concerning manual alignment of words
Let's begin by the cases that will be excluded due to the excessive level of difficulty that they present.
Consider the two sentences:

\[ E_{17} : \text{But the similarities are illusory.} \]
\[ F_{18} : \text{Ces comparaisons sont leurs limites.} \]

Although these sentences correspond to each other in the text that they appeared in, we cannot establish an alignment of their words. We will not study these cases for a few reasons: first, outside of their context it is difficult, even for human readers, to affirm their semantic relation as a translation; secondly, in order to align these sentences, the entire sentences must be considered as an expression, and this is debatable.

How can manual alignments be represented?
We will distinguish the alignment of words and groups of words whose mutual translation is established with the aid of a bilingual dictionary from alignments that are made from a local recomposition based on human "comprehension" of the two sentences. We will use the equal-sign (=) to mark links which come from a bilingual dictionary and the star symbol (*) to mark comprehension correspondences. We will call the first type of correspondence "lexical correspondence" and the second type "contextual correspondence".
The alignments (1-n) (m-1) or (m-n) are characterized by the presence, on the same line, of more than one * or =. Let's give an example:

\[ F_{93} : \text{Une partie seulement de ces vibrations contribuent aux son mis par le clavecin.} \]
\[ E_{120} : \text{Only some of this, vibrational activity contributes to radiating sound of the harpsichord.} \]

\[ F_{93} \quad E_{120} \]
\[ \begin{array}{cccccccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\
\text{Une} & = & \text{partie} & = & \text{s seulement} & = & \text{de} & = & \text{ces} & = & \text{vibrations} & = & \text{contribue} & = & \text{aux} & = & \text{son} & = & \text{mis} & \text{par} & \text{le} & \text{clavecin} & \text{determine} & \text{le} & \text{instrument} \\
\end{array} \]

\[ \begin{array}{cccccccccccccc}
\text{Only} & \text{some} & \text{of} & \text{this} & \text{vibrational} & \text{activity} & \text{contributes} & \text{to} & \text{radiating} & \text{sound} & \text{of} & \text{the} & \text{harpsichord} \\
\end{array} \]
III. Hypothesis
As a basis of our algorithm we find the following hypothesis. Consider two sentences $F$ and $E$ which are translations of each other.
We say that two words $f_i$ and $e_j$, belonging to $F$ and $E$ respectively, correspond to each other if: i) they are translations of each other; ii) they enter into the same dependency relations with their neighbors; iii) they occupy the same positions.

IV. Potential Alignments
Consider the two sentences $F$ and $E$. The potential alignment of words is obtained by comparing each of the words of one sentence with all of those from the second sentence. The comparisons $(f_i, e_j)$ are established with the help of a simple word transfer dictionary and the results are stored in a $m \times n$ matrix (m being the number of words in the French sentence and n in the English). Each element receives a note that is higher if the two words are: i) translations in the dictionary, ii) long, iii) in the same position.

V. Ambiguity, noise and silence
An alignment is 'ambiguous' if more than one solution is produced. Typology of errors (noise, silence): We will call errors of noise those alignments created between words should not be aligned, and errors of silence missing alignments between words which were manually aligned.

VI. The reasons for noise and silence
Noise: At the root of noisy alignments we find the problem of polysemy. When it is not resolved, it causes words to be aligned through sense that are improper in the current context.
Another source of error corresponds to simple errors of alignment: the two words are translations of each other but in the present context they should not be aligned. For example, in the following sentences areas28 was incorrectly aligned with zones17.

F91: Lorsque la table d'harmonie vibre à l'une de ses fréquences de résonance, le paillettes quittent les zones en mouvement ...
E91: When the soundboard vibrates at one of its resonant frequencies, the glitter bounces out of regions that are moving and collects along nodal lines, or in areas where the soundboard is "...

Silence: The main problem is something missing from the dictionary: either the head word is not present, or the correct translation is absent. This is essentially the non-recognition of synonymy that is the problem.

VII. Resolution by Analogic Reasoning
In order to reduce both noise and silence, we use a mechanism based on analogical reasoning. This is based on the following fundamental hypothesis: paradigmatic relations can help determine syntagmatic relations and vice-versa.
Using monolingual dependency relations.
The resolution mechanism can be understood from the following diagram.

\[
\begin{array}{cc}
\text{French} & \text{English} \\
\hline
f_i & R_f & f_j \\
\hline
e_p & R_e & e_q \\
\end{array}
\]

On this figure are represented four words of which two are aligned $(f_i, e_j)$. Syntactic dependencies between two other pairs of words $(f_i, f_j)$ and $(e_p, e_q)$ are also represented. We want to know how valid the alignment between $f_i$ and $e_p$ is.
To answer this, we reason in the following way:
1. On the syntagmatic plane,
   - since $f_i$ is in relation with $f_j$ (the relation $R_f$ being supposed valid),
   - since $e_p$ is in relation with $e_q$ (the relation $R_e$ being valid),
2. on the paradigmatic plane,
   - since $f_i$ is the translation of $e_p$ (supposing the alignment relation $P_2$ is valid),
then we conclude, by analogy, that the alignment relation $P_1$ is also valid, in other words that $f_i$ and $e_p$ are translations of each other in this context. This degree of validity will be stronger as the dependency relations $R_f$ and $R_e$ are close (identical or compatible) and as $P_1$ and $P_2$ get close to identity.
We will call strong resolution one that confirms an existing potential alignment, and weak resolution one that negates an existing alignment or that creates a new alignment.

VIII. Conclusion
The algorithm presented here subdivides into three phases. The first phase is construction: based on lexical proximity, we try to establish all the possible links between the words of the two sentences being aligned. The second phase is one of elimination: using syntactic dependencies we attempt to resolve ambiguous attachments and to undo nonambiguous but erroneous attachments. The third step is again one of construction: we attempt to reduce silence.
We repeat that even human solutions to alignments are subject to wide variations, which shows the difficulty of problem.

Acknowledgements to Hadhemi Achour, Chiraz Ben Othman, Emna Souissi and Gregory Grefenstette.