COMPARATIVE ANALISYS: MACHINE TRANSLATION OF MACROECONOMIC TEXTS

Companies and organizations in the financial sector use translation procedures in direct proportion to their size and, therefore, to the number of texts translated (Arevalillo, 2012; Torres-Hostench, Presas & Cid-Leal, 2016). Specialized translators can contribute by searching for textual patterns to help automate this process as much as possible.

Common evaluation procedures in AT tend to quantify differences between human- and machine-generated texts, to produce metrics which are directly applicable to engine customization, as in the case of BLEU and TER (Molina Baños, 2019).

All types of machine translation (RBMT, SMT or NMT) are based on statistics and are therefore of a quantitative essence (Parra, 2018; Hearne & Way, 2011; Doğru et al., 2017; Bahdanau, Cho & Bengio, 2014). This research adopts a qualitative approach, based on binary/non-binary variations (Pym, 1992), to search for acceptability (Toury, 1980), a burning issue in AT (Rico, 2014; Doherty & Gaspari, 2013; Zaretskaya, 2017).

We therefore delve into the question of differences between artificial and natural texts, as well as the classification, relative prevalence, and repetition patterns of such variations (Scholand, 2013; de la Fuente, 2014, Souto Pico, 2012). Building on BLEU/TER methodologies and adopting a qualitative approach, we classify variations and describe their differential features. We use the MQM metric approach to measuring results (QA; cf. Lommel, Uszkoreit & Burchardt, 2014), which can be applied to both pre-edition (Ramírez Polo, 2012; ISO 17100, par. 4) and post-edition phases. Our aim is to advance in the categorization and delimitation of translational phenomena in macroeconomic texts (García, 2014).

To this end, we follow on the methodology already used in Romana (2009), using viability parameters such as the presence of hypertrophies, heavy sequences, syntactic cohesion and variation in syntactic complexity, as well as prevalence of stylistic trends (Romana, 2009). Research results, such as they are at this stage, are shown below.
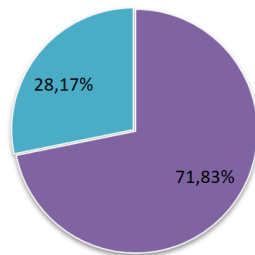
KEYWORDS

Artificial translation, machine translation, macroeconomic translation, quantitative analysis, Levenshtein distance, comparative syntax.
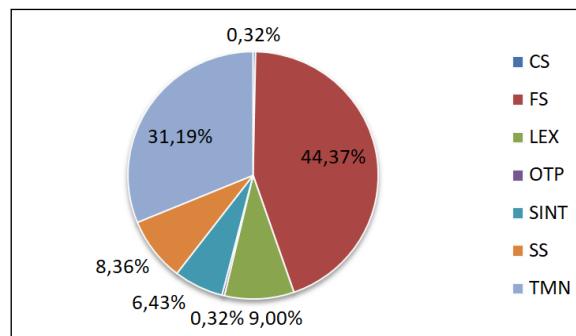
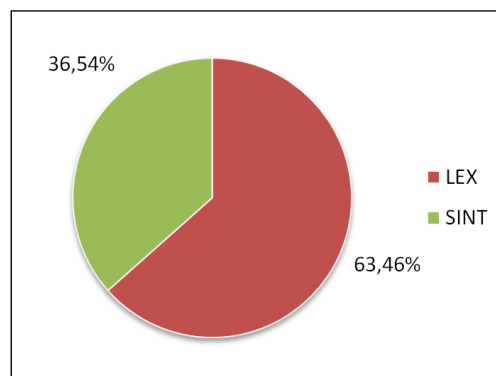| NUMBER OF WORDS (translations; English)	
Total	5466
No change	45
Binary variations	4104
Non-binary variations	1317

GRAPHS

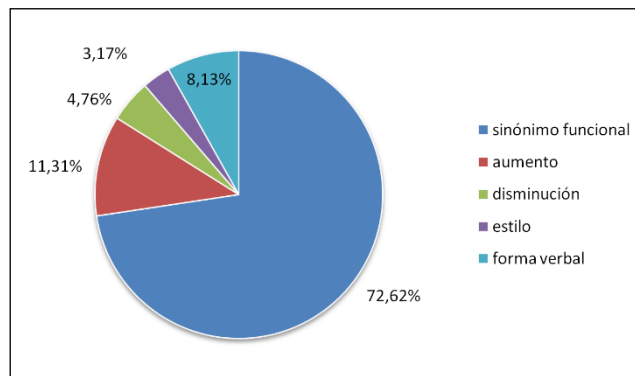- Type of variation (binary 28,17 % / non-binary 71,83 %)

- Distribution by type of binary variation

- Distribution by type of non-binary variation: lexical (63.46 %) and structural (36.54 %)

- Lexical variations

Legend:
- sinónimo funcional — 72,62%
- aumento — 11,31%
- disminución — 4,76%
- estilo — 3,17%
- forma verbal — 8,13%

- Structural variations (syntactic complexity)

Legend:
- mantenimiento — 69,26%
- complicación — 20,27%
- simplificación — 10,47%