Multi-word expressions (MWEs) like *kick the bucket* or *take a walk* stand out as being idiomatic with respect to semantic composition while being restricted to a specific surface realization. In many grammar frameworks, this is dealt with in terms of syntactic ambiguity between an MWE and its literal counterpart – Tree Adjoining Grammar (TAG) being one of them (Abeillé 1995; Abeillé & Schabes 1996). Lichte & Kallmeyer (2016) have recently argued for an approach based on TAG, however, that issues the same syntactic analysis for both and draws the difference solely in the semantics. Hence, on this approach, *kick the bucket* is canonically generated from the TAG elementary trees of *kick*, *the* and *bucket*. Their semantics, on the other hand, is special in that it also carries the meaning of the MWE, thus being systematically ambiguous. Moreover, the MWE part of the semantics is immediately linked to morphological information in order to ensure the right surface realization. An example of this semantic ambiguity approach is shown in Figure 1.

Yet, there are some desiderata for this TAG-analysis of MWEs. Firstly, the MORPH feature is somewhat redundant since the encompassed information is usually also available in the syntactic structure. Secondly, an inference-based extension (following Pulman 1993), which could help to express lexical generalizations more succinctly, seems to come with considerable disadvantages as it presupposes a two-step procedure: first the full literal meaning is composed, and then the MWE meaning is “quasi-inferred” based on the literal meaning. This two-step procedure seems inevitable because, in the TAG-implementation by Lichte & Kallmeyer (2016), “quasi-inferences” on a per-word basis would be hard to control and prone to infinite regress. In my talk, I will address
the desiderata and propose an improved tree grammar model, which I will briefly present in the remainder of this abstract.

While I agree that MWEs should obtain a uniform syntactic analysis, I propose to implement this idea with TUCO (Tree Unification & Constraints), a recently developed tree grammar formalism. TUCO differs from TAG in one crucial respect: In TUCO, the derived tree is composed with only one operation, tree unification, instead of using substitution and adjunction as in TAG. Tree unification is guided by constraints, which are imposed on the derived tree. Those tree constraints are just conditional statements on linked tree and frame descriptions. An example of this sort of constrained composition is shown in Figure 2. What’s more, in TUCO, nodes in the derived tree not only carry sets of morpho-syntactic features, but, unlike in TAG, nodes can be polarized. Polarization eventually makes it possible to see what information has been contributed by elementary trees (i.e. the lexicon), and what is subsequently added by constraints. This distinction becomes substantial when trying to avoid the problem of infinite regress in inference-based approaches: Constraints can now be grounded as to the nodes contributed by elementary trees; in other words, it can be prevented that constraints apply to nodes that are added by constraints, thereby effectively ruling out recursion. As a consequence, it is not only possible to specify the ambiguity between the literal and figurative meaning of *kicked the bucket* as succinctly as in the quasi-inference rule in Figure 3, the quasi-inference rule can also be used at any point of the derivational process, hence allowing for a cognitively more realistic use within incremental processing models.

![Figure 2: TUCO derivation of the literal meaning of kick the bucket: above the brace, the first three columns show the tree unification part; the last column contains one simple constraint. The resulting tree-frame pair is shown below the brace. ①and ② are links connecting syntactic nodes and frame components. • (can unify with anything) and ! (can unify only with • ) are polarization markers.](image-url)
Figure 3: Quasi-inference rule for the figurative meaning of *kick the bucket* in the shape of a TUCO constraint: if the antecedent left of $\Rightarrow$ is fulfilled, wrap an $\mathcal{I}$-term around the semantic part of the antecedent and the consequent in the following way: $\tau \land a \Rightarrow b = \tau \land a \Rightarrow (a \mathcal{I} b)$ with $\tau$ being a tree description and $a$ and $b$ frame descriptions. $\langle\text{kick}\rangle$ marks the baseform of kicked.


