Backward control in Modern Standard Arabic

Tali Arad Greshler and Nurit Melnik

1 Introduction

Subjunctive complement clauses in Modern Standard Arabic (MSA) are used in contexts where English (and other languages) uses the infinitives. However, unlike English infinitivals, subjunctives in MSA exhibit agreement. Moreover, they alternate between control interpretations, where the matrix subject and the embedded subject share reference, and no-control interpretations, where they have disjoint reference. A corpus-based investigation conducted by Arad Greshler et al. (2016) (henceforth AHMW) revealed that there are no obligatory control (OC) predicates in MSA. Consequently, they propose a uniform *pro*-drop based analysis of control and no-control. The analysis accounts for all patterns, except one – backward control – which exhibits a surprising agreement pattern. AHMW explain the discrepancy by suggesting that this construction may involve extra-grammatical factors.

This study proposes an alternative account of the backward control construction, which builds on new corpus findings regarding the types of predicates which are licensed in this construction. We propose that these predicates can optionally form complex predicates with the embedded subjunctives. When this occurs, the complex predicate exhibits the regular agreement patterns associated with VSO and SVO clauses in MSA.

2 Background

MSA is a *pro*-drop language whose unmarked word order is VSO, yet SVO order is also available. While the two word orders are possible, each is associated with a different agreement pattern. Post-verbal subjects trigger partial agreement, which only involves gender, while the number feature is invariably singular. Pre-verbal subjects trigger full agreement on the verb.

MSA subjunctive clauses are preceded by a the particle *ʔan* and are obligatorily verb-initial. They typically resemble control constructions in English (and other languages), where an unexpressed subject is controlled by a matrix argument. However, the agreement marking on the subjunctive verb reveals the agreement properties of the intended subject. In (1a) the subjunctive *yaktuba* '*write*' exhibits 3SM agreement. Consequently, the understood embedded subject can be construed as the matrix subject (*control*) or as a different singular–masculine referent (*no-control*). The control and no-control interpretations are also possible in the backward pattern illustrated in (1b), where the subject appears in the embedded clause. In addition, since there is no OC, embedded subjunctives may exhibit agreement properties distinct from the matrix predicate. In (2) the embedded subject is optional but control is impossible due to the agreement mismatch (matrix 3SM and embedded 3SF).

(1) a. *ħaawala muħammad-un [ʔan yaktuba maqaal-an.]*
*tried.*3SM *Muhammad-*NOM(M) *AN write.*3SM.SBJ *article.*ACC
'Muhammad tried to write an article.'
'Muhammad$_i$ tried that he$_j$ would write an article.'

b. *ħaawala [ʔan yaktuba muħammad-un maqaal-an.]*
*tried.*3SM(M) *AN write.*3SM.SBJ *Muhammad-*NOM(M) *article.*ACC
'Muhammad tried to write an article.'
'He tried that Muhammad would write an article.'

(2) *ħaawala muħammad-un [ʔan taktuba (hind-un) maqaal-an.]*
*tried.*3SM *Muhammad-*NOM(M) *AN write.*3SF-SBJ *(Hind-*NOM(F)*) article.*ACC
'Muhammad tried that Hind/she would write an article.'

3 Subjunctive reference patterns in MSA

AHMW conducted a corpus-based investigation with the goal of finding whether all *ʔan*-clause selecting predicates allow for both control and no-control between the two subjects, or whether there are OC predicates. The corpus that was used is the 115-million-token sample of the *arTenTen* corpus of Arabic (Arts et al., 2014).

The corpus investigation led AHMW to conclude that there are no OC predicates in MSA. They found evidence for control and no-control with various types of predicates: volitionals, implicatives, manipulatives, modals, and aspectuals. These findings echo Habib (2009), who assumes that there are no "real" control predicates in MSA. They do constitute, however, counterexamples to the generalization made by Landau (2013, p.106), who predicts that "[t]here cannot be a language where modal, aspectual and implicative verbs or evaluative adjectives allow an uncontrolled complement subject", provided that the embedded predicate exhibit morphological agreement.

Under the assumption that there is no OC in MSA, AHMW argue for one structure for all cases, namely, a no-control structure. Constructions with *ʔan* complement clauses are simply structures with two independent subjects. The omission of a subject in either clause is due to the *pro*-drop property of MSA; each of the clauses, the matrix clause and the embedded clause, can either have an overt subject or a *pro*-dropped subject. There are no constraints on the agreement relations between the two predicates, and therefore they do not need to match. What resembles subject control is in actuality co-indexation at the semantico-pragmatic level.

One pattern proved problematic for this analysis. The simple example of the backward pattern in (1b) masks a more complex agreement pattern which is only discernable with plural human subjects, for which agreement varies depending on the position of the subject relative to the verb. AHMW find that when the embedded subjects are both human and plural the matrix verb exhibits partial agreement (i.e., only in gender and person) with the subject (3).

(3) *ħaawalat [ʔan taktuba$_i$ l-banaat-u$_i$ maqaal-an].*
*tried.*3SF *AN write.*3SF.SBJ *the-girls*-NOM *article*-ACC
'The girls tried to write an article.'

This is unexpected under the *pro*-drop analysis. *Pro* subjects are assumed to trigger full agreement on their predicates. If so then it is not clear how a 3SF *pro* can co-refers with the plural embedded subject.

AHMW conclude that there is no evidence for the existence of OC predicates in MSA. A one-structure *pro*-drop analysis accounts for most of the data, with the exception of the agreement pattern attested in the backward construction (3). They suggest that the use of partial agreement in this pattern is motivated by analogy to the partial subject–verb agreement found in simple VSO clauses, and that the integration of this construction into the theory requires some additional assumptions, which may involve extra-grammatical factors, possibly related to the non-native status of MSA.[1]

4 The distribution of backward control

The goal of this study is to propose an alternative account of the backward control construction illustrated in (3). It begins by conducting more focused corpus investigations of the backward pattern. First, we ask whether it is indeed the case that there are no instances of full agreement when the subject is expressed in the embedded clause. Moreover, we extend the range of predicates examined by AHMW to investigate whether all predicates are compatible with backward control.

[1]MSA is the literary standard of the Arab world, but it is acquired in school. The mother tongue of its speakers is some regional dialect of Arabic.

A corpus investigation limited to cases with plural animate subjects revealed instances of backward control with volitions, implicatives, modals and aspectuals. No instances of backward control were found with the following predicates : *qarrara* '*decide*', *xaʃiya* '*fear*', *rafadˤa* '*refuse*', *tarradada* '*hesitate*', and *ʔiqtaraħa* '*propose*'.

With regards to agreement, contrary to AHMW, we found instances of full agreement on the matrix predicate. However, unlike a similar raising construction discussed by Wurmbrand & Haddad (2016), whose matrix predicates alternate between full and partial agreement, the difference in the agreement marking was found to affect the interpretation of the two sentences. When the embedded subject is plural and human and the matrix predicate exhibits partial agreement with it the sentence is ambiguous (4a). The unexpressed matrix subject can be construed as the embedded subject (control) or as a singular-feminine referent (no-control). When the matrix predicate is plural, there is only one no-control interpretation (4b).

(4) a. *ħaawalat*$_{i/j}$ *[ʔan taktuba*$_i$ *l-banaat-u*$_i$ *maqaal-an].*
*tried.*3SF *AN write.*3SF.SBJ *the-girls*-NOM *article*-ACC
'The girls tried to write an article.' / 'She$_j$ tried that the girls$_i$ would write an article.'

b. *ħaawalna*$_{*i/j}$ *[ʔan taktuba*$_i$ *l-banaat-u*$_i$ *maqaal-an].*
*tried.*3PF *AN write.*3SF.SBJ *the-girls*-NOM *article*-ACC
'They$_j$ tried that the girls$_i$ would write an article.' / Not: 'The girls tried to write an article.'

This asymmetry does not occur in the forward pattern (1a); the two interpretations (control/no-control) are always possible and the embedded predicate exhibits full agreement with its construed subject.

A similar phenomenon is found in Modern Greek (MG), a language which shares a number of syntactic properties with MSA. Subjunctive complement clauses in MG fall into two categories: controlled-subjunctives (C-subjunctives), which enforce control between the matrix and embedded subject, and free-subjunctives (F-subjunctives), which, like in MSA, allow for both control and no-control. Also similarly to MSA, in both types of constructions the subject can be expressed either in the matrix clause or the embedded clause, yet the backward pattern with F-subjunctives is more restricted .

With C-subjunctives a control interpretation is the only option regardless of the position of the subject. With F-subjunctives, on the other hand, the forward pattern in (5a) is ambiguous between control and no-control, but in the backward pattern (5b) the embedded subject cannot be controlled by the matrix subject (Alexiadou et al., 2010, ex. 39). This is similar to the MSA data in (4b).

(5) a. *o Janis*$_i$ *elpizi [na fai pro*$_{i/j}$ *to tiri]*
John-NOM *hopes subj eats pro the cheese*
'John$_i$ hopes that he$_{i/j}$ will eat the cheese.'

b. *pro*$_{*i/j}$ *elpizi [na fai o Janis*$_i$ *to tiri]*
pro hopes subj eats John-NOM *the cheese*
'He hopes that John will eat the cheese.'

Alexiadou et al. (2010) propose a *pro*-drop analysis for F-subjunctives, similar in spirit to the one proposed by AHMW. Consequently, they attribute the impossibility of coreference in (5b) to Principle C. The embedded referential subject, *Janis*, cannot be bound by the matrix *pro* subject. The fact that there is no Principle C effect in the case of C-subjunctives is taken by Alexiadou et al. (2010) as evidence that control with these predicates does not involve a *pro*-dropped subject.

The similarity between MSA and MG is even greater when the types of predicates which are licensed by the different constructions are considered. Landau (2004) argues that the predicates which select C-subjunctives in MG belong to a category of predicates which cross-linguistically select *semantically untensed* complements, and include the implicatives, aspectuals, modals, and evaluative adjectives. Predicates which select F-subjunctives, on the other hand, are those which select *semantically tensed*

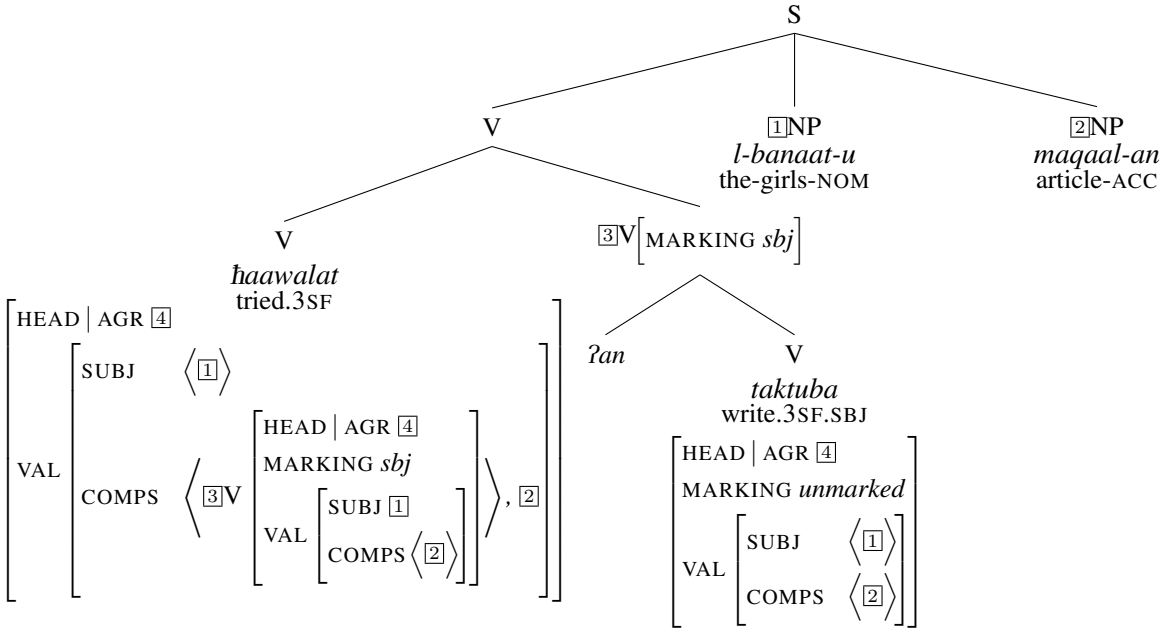
complements (factives, propositional, desiderative, and interrogatives). Our corpus investigations reveal that the predicates which are licensed in backward control in MSA belong to the same category as those which select C-subjunctives in MG. Thus, although AHMW found that they are not OC predicates, contrary to Landau's (2013) prediction, the association between this category and backward control in MSA cannot be coincidental.

5 Complex predicates and control

Complex predicates is a term that is used to describe a situation whereby two (or more) predicates function as a unit in a monoclausal structure. This is also referred to as 'restructuring' in the context of infinitival complements in Germanic and Romance languages (Wurmbrand, 2001). Roussou (2009) proposes a conceptually similar analysis for subjunctive complements in MG which she refers to as 'clause-union'. She argues that since C-subjunctives lack semantic tense they do not constitute an independent event. Consequently, they trigger clause-union with their selecting predicate and "event composition leads to composition of argument structure as well" (Roussou, 2009, p.1827). F-subjunctives, on the other hand, do not trigger clause-union. A similar proposal is made by Grano (2015).

A number of properties exhibited by backward control in MSA motivate a complex-predicate analysis. First, the predicates which are licensed in this construction belong to the same category as those which select C-subjunctives. Similarly to the MG predicates, the embedded clauses of these predicates cannot be temporally modified independently from the matrix clause. Having only one tense associated with a construction suggests a monoclausal structure. Second, there are strict adjacency conditions with respect to the linear position of the selecting predicate, *ʔan*, and the subjunctive. Finally, by proposing that the backward control construction has a monoclausal structure the partial agreement on the matrix predicate (as well as the embedded predicate) is expected since the two predicates precede their (shared) subject.

Our proposal augments AHMW's analysis of no-control and forward control. Following Alexiadou et al. (2010), we explain the unavailability of the coreference reading in the backward pattern (4b) by invoking Principle C. Moreover, we follow previous HPSG analyses of complex predicates that have been used to account for phenomena in diverse languages (e.g., Hinrichs & Nakazawa, 1990; Abeillé et al., 1998; Monachesi, 1998; Müller, 2002, among others) and propose a similar analysis for backward control in MSA. A sketch of our proposed analysis of (3) is presented below.



In a nutshell, we propose that *ʔan* and the subjunctive combine in a head-marker structure. The matrix predicate selects as its complement a marked subjunctive with matching agreement properties, and inherits all its arguments. Moreover, the referential index of the shared subject is structure-shared with the values of the semantic roles of the relations denoted by each of the predicates, as is expected in a control construction (not shown in the sketch due to lack of space).

In the full version of the paper we show that AHMW's analysis together with the current proposal account for all the patterns attested with *ʔan* complement clauses. This analysis is given support by cross-linguistic findings regarding parallel predicate types and their syntactic behavior. Consequently, what AHMW consider a surprising pattern can be explained with known syntactic mechanisms.

References

Abeillé, Anne, Danièle Godard & Ivan A. Sag. 1998. Two kinds of composition in French complex predicates. In E. Hinrichs, A. Kathol & T. Nakazawa (eds.), *Complex predicates in nonderivational syntax*, vol. 30 Syntax and Semantics, San Diego: Academic Press.

Alexiadou, Artemis, Elena Anagnostopoulou, Gianina Iordachioaia & Mihaela Marchis. 2010. No objections to backward control. In M. Polinsky & N. Hornstein (eds.), *Movement theory of control*, 89–117. Amsterdam: John Benjamins.

Arad Greshler, Tali, Livnat Herzig Sheinfux, Nurit Melnik & Shuly Wintner. 2016. Reference patterns in subjunctive complement clauses of modern standard arabic. In D. Arnold, M. Butt, B. Crysmann, T. H. King & S. Müller (eds.), *Proceedings of the joint 2016 conference on head-driven phrase structure grammar and lexical functional grammar, polish academy of sciences, warsaw, poland*, 4–22. Stanford, CA: CSLI Publications.

Arts, Tressy, Yonatan Belinkov, Nizar Habash, Adam Kilgarriff & Vit Suchomel. 2014. arTenTen: Arabic corpus and word sketches. *Journal of King Saud University - Computer and Information Sciences* 26(4). 357 – 371. doi:http://dx.doi.org/10.1016/j.jksuci.2014.06.009.

Grano, Thomas. 2015. *Control and restructuring*. Oxford: Oxford University Press. (Oxford Studies in Theoretical Linguistics series).

Habib, Rania. 2009. The syntax of the Standard Arabic *ʔan* and *ʔanna*. In K. K. Grohmann & P. Panagiotidis (eds.), *Selected papers from the 2006 Cyprus syntaxfest*, 159–194.

Hinrichs, Erhard & Tsuneko Nakazawa. 1990. Subcategorization and vp structure in german. In S. Hughes & J. Salmons (eds.), *Proceedings of the third symposium on germanic linguistics*, Amsterdam: Benjamins.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22(4). 811–877.

Landau, Idan. 2013. *Control in generative grammar: A research companion*. Cambridge University Press.

Monachesi, Paola. 1998. Italian restructuring verbs: A lexical analysis. In E. Hinrichs, A. Kathol & T. Nakazawa (eds.), *Complex predicates in nonderivational syntax*, vol. 30 Syntax and Semantics, 313–368. San Diego: Academic Press.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford: CSLI Publications.

Roussou, Anna. 2009. In the mood for control. *Lingua* 119(12). 1811–1836.

Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.

Wurmbrand, Susi & Youssef A Haddad. 2016. Cyclic spell-out derived agreement in Arabic raising constructions. In Y. A. Haddad & E. Potsdam (eds.), *Perspectives on Arabic linguistics XXVIII: Papers from the annual symposium on Arabic linguistics, Gainesville, Florida, 2014*, 193–228. Amsterdam/Philadelphia: John Benjamins Publishing Company.