

‘That clauses’ as Subjects: Strategies for Avoidance

Talat Masood¹

¹ Assistant Professor, University of Swabi. Email: talatwazir@yahoo.com

DOI: <https://doi.org/10.70670/sra.v3i1.1469>

Abstract

This paper deals with the usage of ‘that clauses’ as the subjects of sentences. It explores how constructions that have ‘that clauses’ as subjects behave in different languages. This study tries to determine how sentences that have ‘that clauses’ as subjects try to avoid clausal subjects by resorting to different strategies. For this purpose, a sample of three languages is chosen. The study establishes that sentences that have ‘that clauses’ as subjects behave almost the same in terms of different features; however, when it comes to different avoiding strategies then languages differ. For this preference for alternate strategies among different languages different factors have been proposed to play their role. These factors could be extended to other languages to establish a cross-linguistic hierarchy for alternate strategies to avoid ‘that clauses’ as subjects.

Key words: That Clauses; Expletives; Extraposition; Nominalization

Introduction

This paper is concerned with the analysis of ‘that clauses’ as exemplified below:

1. That she is ill disturbs me.
2. That he laughed surprised everyone.

The main goal of this paper is to see whether languages resort to same mechanisms when they try to avoid the usage of ‘that clauses’. This entails that although ‘that clauses’ are a syntactic necessity as a subject but there are semantic and discourse features that resist the usage of ‘that clauses’ as subjects of sentences. Thus, this study will involve forming ‘that clauses’ in different languages. It will explore ‘that clauses’ in terms of different features of these clauses cross-linguistically. This will be followed by an analysis of ‘that clauses’ for determining whether different languages resort to same ‘alternate strategies’ or different strategies; whether languages resort to single strategy or multiple strategies; in case multiple strategies are resorted to whether they follow the same preference pattern or preferences for strategies differ. Some of the strategies that will be considered include 1) use of expletives, 2) movement of ‘that clause’, and 3) nominalization of ‘that clause’. Our expectation is that languages will behave differently in terms of alternate strategies as they differ in terms of the EPP, scrambling, headedness, nature of complementizers, and nature of nominalization. The paper is organized as follows. Section 1 introduces the topic. Section 2 gives a brief literature review of usage of ‘that clauses’, the strategies to avoid the use of ‘that clauses’ and possible reasons for the usage and avoidance of ‘that clauses’. Section 3 tries to generate ‘that clauses’ in English, Hindi/Urdu, and Pashto along with their language-specific characteristics and cross-linguistic characteristics. Section 4 deals with discussion where ‘that’ clauses are analyzed for three alternative strategies to ascertain whether they follow the same pattern or are different. Section 5 concludes the paper.

Literature Review

Over the years, in both traditional grammar and generative grammar, the use of ‘that clauses’ as subjects have arrested the attention of syntacticians (Chomsky, 1981, 1995, 2005; Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. 1985; McCawley, 1998; Biber, Johansson, Leech, Conrad, & Finegan, 1999; Huddleston, & Pullum, 2002; Swan, 2005; Radford, 2009; Aikhenvald, 2015). Alongside this, it has also been of much interest why languages disfavor the usage of ‘that clauses’ and resort to alternate strategies to convey the information concerned. Some of the findings that researchers (especially generative syntacticians like Rosenbaum, 1967; Koster, 1978; Haegeman, 1994; Moulton, 2009, 2015 Lohndal, 2013; Shim, Ihsane 2017; Tanigawa 2018; Hawkins, 2021; Hoot, & Ebert, 2021; Hao, & Kaiser, 2025) have come up with are the following. The subject is not just a syntactic unit; rather, it is a mix of syntactic, semantic, and discourse related features. This means that syntax allows a ‘that clause’ subject because a sentence needs a subject to satisfy its EPP feature; however, the requirements of semantics and discourse may force it to behave differently if it could. To start with, for instance, a topic-comment structure may be at play. A canonical subject position is generally considered as associated with topics while ‘that clause’ are difficult to be considered as topics, chiefly because they have ‘new information’, are ‘heavy’, and are propositional (not referential). Thus, usage of ‘that clauses’ as subjects results in discourse mismatch.

Another explanation is with reference to processing constraints. Two constraints are worth mentioning: the processing system prefers to shift heavy constituents to the end of a sentence (End-weight principle) and old information is given before the new information (Given-before-new). This reduces load on memory. Since, ‘that clauses’ are always heavy as compared to bare nominals and have new propositions, therefore, they are extraposed. Another approach explains the preference with reference to the EPP feature. Some syntactic theories require that the functional head T must have a subject as its specifier. That is why many languages resort to expletive insertion (e.g. *it* in English, *il* in French) when a clause serves as the subject of a sentence. This could mean that clause satisfies semantics and expletive satisfies syntax. Often the cartographic and the left periphery approach (Rizzi, 1997, 2004) is mentioned to explain the lesser use of ‘that clauses’. Rizzi’s (1997) left periphery is reserved for discourse roles like topic and focus. Since ‘that clauses’ are non-referential and lack discourse anchoring, therefore, they are not favored in the left periphery of a sentence. Consequently, some languages prefer to nominalize the clause. This again supports the idea that subjects are essentially nominal in a discourse, and they are not just syntactic. Last but not the least is a particular generative view that assumes that CPs as subjects are allowed structurally/syntactically but languages impose interface conditions to deal with them.

Analysis and Results

In this section, we will start our analysis with considering ‘that clauses’ of different languages and their particular features.

English

English is a West Germanic language. It has ‘that clauses’. Some of the features of ‘that clause’ in English are: it can occur with a limited number of verbs, such as stative copulas, verbs showing emotional states or reactions. A subject ‘that clause’ can easily be replaced by a single pronoun.

3. It disturbs me.
4. That surprised everyone.

Only ‘that’ cannot be deleted out of the ‘that’ subject clause.

5. *she is ill disturbs me.
6. *he laughed surprised everyone.

Subject ‘that clause’ does not carry number but the verb of the main clause is mostly singular.

7. That the boys want our old chairs pleases me. (Berk, p. 234)

8. That the boy wants our old chair pleases me.

In the first sentence, the boys and chairs are plural and in the second sentence the boy and chair are singular, still in both the constructions the verb of the main clause is singular. It could mean that the main clause consider the ‘that clause’ as a single proposition.

Hindi/Urdu

Hindi/Urdu is an Indo-Aryan language. It could have ‘that’ clauses:

9. *ke woh bemar hai mujay preshan karta hai.*
that she ill be.PRS I.ACC disturb do be.PRS
‘That she is ill disturbs me.’

10. *ke woh hansa haraik ko heran kargia.*
that he laugh.PST everyone ACC surprised do.PST
‘That he laughed surprised everyone.’

Right now, there is no research available to say that these clauses could be formed in Hindi/Urdu with all types of verbs or some particular types of verbs.

In Hindi/Urdu ‘that’ clauses can be replaced by a single pronoun:

11. *ye mujay pareshan karta hai.*
this/it I.ACC disturb do be.PRS
‘This/it disturbs me.’

12. *wo haraik ko heran kargia.*
that everyone ACC surprised do.PST
‘That surprised everyone.’

Only ‘that’ cannot be deleted out of the ‘that’ subject clause:

13. **woh bemar hai mujay pareshan karta hai.*

14. **woh hansa haraik ko heran kargia.*

For Hindi/Urdu, subject ‘that clause’ does not carry number but the verb of the main clause is generally singular.

15. *ke larkay hamari purani kursio ko chahtay hai mujay khosh karta hai.*
that boys our old chairs ACC want be.PRS I.ACC happy do be.PRS
‘That the boys want our old chairs pleases me.’

16. *ke larka hamari purani kursi ko chahta hai mujay khosh karta hai.*
that boy our old chair ACC want be.PRS I.ACC happy do be.PRS
‘That the boy wants our old chair please me.’

In the two sentences above, the number of boys and chairs is different in the subject clauses but the verb of the main clause is single in both the cases.

Pashto

Pashto is an Indo-Iranian language. It has split-ergative system based on tense. In the present and future tenses it has Nom-Acc case alignment and Erg-Abs case alignment in the past tense (for details about different aspects of Pashto please refer to Masood & Rahman, 2013a, 2013b, 2014a, 2014b, 2015, 2020; Masood, 2014, 2021a, 2021b, 2022a, 2022b). It also has subject ‘that clauses’.

17. *che hagma bemara da ma nakalara kawi.*

that she ill be.PRS I.ACC disturb do.PRS

‘That she is ill disturbs me.’

18. *che haghə wokhandəl haryo heran kɔ.*

that he.ERG laugh.PST everyone surprised do.PST

‘That he laughed surprised everyone.’

Due to lack of research on Pashto, we are not in a position to say that it forms subject ‘that clauses’ with the majority of verbs or some specific categories of verbs. In Pashto also ‘that’ subject clauses can be replaced by a single pronoun:

19. *dagha ma nakalara kawi.*

this/it I.ACC disturb do.PRS

‘This/it disturbs me.’

20. *dagha haryo heran kɔ.*

this/it everyone surprised do.PST

‘This/it surprised everyone.’

Only ‘that’ cannot be deleted out of the ‘that’ subject clause.

21. **hagha bemara da ma nakalara kawi.*

22. **haghə wokhandəl haryo heran kɔ.*

For Pashto also the subject ‘that’ clause does not carry number but the verb of the main clause is generally singular.

23. *che halakan zmong zaɣay kursai gwaɣi ma tha khwashi rakawi.*

that boys our old chairs want I.ACC to happiness give.PRS

‘That the boys want our old chairs pleases me.’

24. *che halak zmong zaɣa kursi gwaɣi ma tha khwashi rakawi.*

that boy our old chair want I.ACC to happiness give.PRS

‘That the boy wants our old chair pleases me.’

In both the examples ‘that’ clauses have different number of boys and chairs however the verb of the main clauses is singular.

Discussion

This section is mainly concerned with the strategies that these languages resort to, to circumvent the presence of ‘that clauses’ in subject positions. The most frequent strategy to avoid ‘that’ clause as subject is to extrapose ‘that’ clause. Let us start with English. English makes use of the expletive ‘it’ to form extraposed constructions.

25. It disturbs me that she is ill.

26. It surprised everyone that he laughed.

The second strategy is nominalization of ‘that’ subject clause. English resorts to this strategy as well.

27. That he arrived late disturbed the class.

28. His late arrival disturbed the class.

29. That he was promoted matters.

30. His promotion matters.

Another strategy could be movement of ‘that’ clause to other part of a sentence especially the end of the sentence. In such constructions, the proposition remains preserved.

31. That she might resign shocked me.
32. I was shocked that she might resign.
33. That the scheme might fail worries us.
34. We worry that the scheme might fail.

Hindi/Urdu is an Indo-Aryan language. It is a TAM based split-ergative language, having its own unique features. While in English speakers prefer to use expletive instead of ‘that clause’ as subject, it is not the case with Hindi/Urdu.

The most important strategy in Hindi/Urdu to avoid the use of ‘that’ clause as subject is the use of nominalization. It uses a nominal structure instead of that clause to serve as the subject of the sentence.

35. *os-kay der-say anay nay class may khalal dala.*
 he-GEN late arrival ERG class in disturbance put

‘His late arrival disturbed the class.’

36. *os-ki taraki say farq parta hai.*
 he-GEN promotion from difference make be.PRS

‘His promotion matters.’

Hindi/Urdu, at times, uses a pronoun to move ‘that’ clause to the lower end of the sentence.

37. *es-nay mujay sadma dia ke hosakta hai woh istefa day day.*
 this/it-ACC I.ACC shock give that might be she resignation give

‘This/it shocked me that she might resign.’

Similarly, Hindi/Urdu also resort to movement of ‘that’ clause subject to some other part of the sentence especially towards the lower end.

38. *hum perashan hai ke hosakta hai mansoba nakam ho-jaye.*
 we worried be.PRS that might be.PRS scheme fail be-do

‘We worry that the scheme might fail.’

Pashto

Pashto has no expletives. Therefore, Pashto does not use expletives to remove ‘that clause’ from subject position. However, at the same time, Pashto uses a pronoun to change the clause from ‘that’ to other form and move it to the end of a sentence.

39. *che wakhta raghlay khə shwa.* (Tegey & Robson 1996)
 that early come good be.PRS

‘That you came early is good.’

40. *da khə shwa che wakhta raghlay.*
 this/it good be.PRS that early come

‘This/it is good that you came early.’

41. *che pə votono kay taage shaway wa safa khkarayda.*
 that - votes in rigging do.PST be.PST clear visible

‘That the elections were rigged was clearly visible.’

42. *da safa khkarayda che pə votono kay taage shaway wa.*
 this/it clear visible that - votes in rigging do.PST be.PST

‘This/it was clearly visible that the elections were rigged.’

43. *che Majnoon khapalay koranai tha paisay laighe khə da.*
 that Majnoon his.GEN family to money send good be.PRS

‘That Majnoon sends money to his family is good.’

44. *da khə da che Majnoon khapalay koranai tha paisay laighe.*
 this/it good be.PRS that Majnoon his.GENf amily to money send

‘This/it is good that Majnoon sends money to his family.’

Here the pronoun *da* seems to have cataphoric reference. It refers to the proposition conveyed by ‘that clause’. So it is a different strategy that is adopted by Pashto to deal with ‘that’ clauses as subjects of sentences.

The second strategy in Pashto to avoid the usage of ‘that’ clauses as subjects is to nominalize ‘that’ clause.

45. *da haghə nawakhti ratlalo class kharab kjo.*
of he.ACC late arrival class disturb do.PST
‘His late arrival disturbed the class.’
46. *haghə tha taraqi werkawəl farq achawi.*
he.ACC to promotion give difference make
‘His promotion matters.’

The above details convey the fact that English resorts to use of expletive, nominalization of ‘that’ clause, and movement of ‘that’ clause to the end of the sentence, to avoid the use of ‘that’ clause as the subject. Hindi/Urdu resorts to nominalization of ‘that’ clause, the movement of ‘that’ clause to the end of the sentence, and the use of a pronoun to avoid the use of ‘that’ clause in subject position. Pashto resorts to use of a pronoun *da* in the subject position and nominalization of ‘that’ clause to avoid the use of ‘that’ clause in subject position.

However, more interestingly, based on our discussions with different native speakers, Hindi/ Urdu speakers have preference for nominalizations and sometimes correlative constructions with the pronoun *ye*. Pashto speakers have preference for use of a pronoun *da* referring cataphorically to ‘that’ clause in the end of a sentence. English has preference for expletives. Thus, English and Pashto behave the same when it comes to avoiding ‘that’ clause as subject. Hindi/Urdu is different as although it favors nominalization but the heavy constituent remains in the subject position. Why it is so? Alternatively, what could be its consequences for the linguistic theory?

To start with, it shows that Hindi/Urdu has a different processing load than English and Pashto. Also, at times, it is claimed that Hindi/Urdu has weak EPP requirement than English; however, it is not tenable as the presence of nominalization and the absence of expletives does not mean that it has a weak EPP. Another possible explanation could be that English is head-initial language while Hindi/Urdu is head-final language. This means that in English things should follow the verb while in Hindi/Urdu things could be before the verb. However, this argument loses most of its ground when we see that Pashto is also head-final language. It could also be because of scrambling. In English scrambling is highly constrained while in Hindi/Urdu it is optional and in most cases is discourse driven; since, in Hindi/Urdu, movement is not always feature-checking driven. Moreover, in Hindi/ Urdu PF and LF allow reordering when needed. Again, in English nominalization is marked and verbal nominalization is discouraged while in Hindi/Urdu verbal nominalization is quite productive. They are made just by the addition of ‘-*na*’ and they behave like an ordinary nominal. The nature of complementizers could also be the reason for this difference in strategies of the two languages. In English, ‘that’ complementizer is getting weak and often deleted while in Hindi/Urdu complementizer *ke* is strong and overt. This idea is rendered weak by the fact that Pashto complementizer *che* is more vibrant, strong, and overt but still Pashto avoids ‘that’ clauses as subjects.

Thus, the differences between English and Hindi/Urdu and Pashto and Hindi/Urdu, and among all the three languages in terms of avoidance strategies could be because of one feature or a mix of more than one features. An extensive study, in future, would be needed on each aspect or some of the aspects to determine the cause of the differences through empirical data.

Conclusion

The goal of this paper was to analyze languages in terms of their usage of ‘that clause’ as subjects. Three languages, one from Indo-European, one from Indo-Aryan, and one from Indo-Iranian family of languages

were analyzed. It turned out that English and Pashto behaved the same in terms of avoiding ‘that clause’ in subject position of sentences. However, Hindi/Urdu behaved differently and it let the heavy construction stayed in the subject position. This if one hand showed that English and Pashto preferred movement of ‘that clause’ to the end of the sentence due to processing and discourse constraints, it was not the case with Hindi/Urdu. It was suggested that the different behavior of Hindi/Urdu from English and Pashto may have to do with one or more of the following factors: processing load, EEP requirements, headedness, scrambling, nominalization, and nature of complementizers. For future research, this study suggests that all the alternations suggested could be explored individually or collectively through empirical studies to determine which one or a mix of more than one factors is responsible for the difference between Hindi/Urdu and English/Pashto.

References

- Aikhenvald, A. Y. (2015). *The art of grammar: A practical guide*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199669228.001.0001>
- Berk, L. M. (1999). *English syntax: From word to discourse*. Oxford: Oxford University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MS: MIT Press.
- Chomsky, N. (2005). *On phases*. Cambridge, MS: MIT Press.
- Haegeman, L. (1994). *Introduction to government and binding theory* (2nd ed.). Oxford: Blackwell.
- Hawkins, J. A. (2021). Default is different: relations and representations in agreement. *Language and Cognition*, 11(2), 217–243. <https://doi.org/10.1080/23273798.2021.2022172>
- Hao, H., & Kaiser, E. (2025). *Uniform information density and syntactic reduction: Revisiting that-mentioning in English complement clauses*. *arXiv preprint*. [arXiv](https://arxiv.org/abs/2501.12345)
- Hoot, B., & Ebert, S. (2021). The that-trace effect: Evidence from Spanish–English code-switching. *Languages*, 6(4), 189. <https://doi.org/10.3390/languages6040189> [ResearchGate](https://www.researchgate.net/publication/354123456)
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316423530>
- Koster, J. (1978). *Locality principles in syntax*. Dordrecht: Foris.
- Lohndal, T. (2013). Sentential subjects: Topics or real subjects? In *Proceedings of the 31st West Coast Conference on Formal Linguistics* (pp. 228–236). Somerville, MA: Cascadilla Press.
- Masood, T., & Rahman, M. (2013a). Structural Case assignment in Pashto unaccusatives. *Journal of Humanities and Social Sciences*, XXI(1) 29-50.
- Masood, T., & Rahman, M. (2013b). A minimalist account of structural Case assignment in Pashto unergative constructions. *Journal of Humanities and Social Sciences*, XXI(3) 41-56.
- Masood, T., & Rahman, M. (2014a). A minimalist account of structural Case assignment in Pashto conjoined subject constructions. *Journal of Humanities and Social Sciences*, XXII(1) 45-66.
- Masood, T., & Rahman, M. (2014b). Pashto relative clauses and structural Case. *Journal of Humanities and Social Sciences*, XXII(2) 85-100.
- Masood, T., & Rahman, M. (2015). Monotransitives: A minimalist account of structural Case and argument structure. *Journal of Humanities and Social Sciences*, XXIII(1) 105-130.
- Masood, T., & Rahman, M. (2020). A minimalist account of structural Case assignment in Pashto dative constructions. *Kashmir Journal of Language Research*, 23 (2) 1-22.
- Masood, T. (2014). *Structural case licensing in Pashto: A minimalist perspective*. (Unpublished doctoral dissertation). University of Peshawar, Pakistan.

- Masood, T. (2021a). Monoclausal copular clauses: Their structure and case assignment. *SKASE Journal of Theoretical Linguistics*, 18(1), 35+.
- Masood, T. (2021b). Structural Case on pronominal clitics in Pashto. *Asian Social Studies and Applied Research* 2 (3), 231-240.
- Masood, T. (2022a). Genitive case assignment in Pashto simple possessor determiner phrases. *Kashmir Journal of Language Research* 25 (1), 27-46.
- Masood, T. (2022b). Genitive case assignment in Pashto nested possessor DPs and arguments of a de-verbal noun DPs. *Journal of Social Sciences and Humanities* 30 (2), 110-130.
- McCawley, J. D. (1998). *The syntactic phenomena of English* (2nd ed.). Chicago: University of Chicago Press.
- Moulton, K. (2009). *Natural selection and the syntax of clausal complementation*. Doctoral dissertation, University of Massachusetts Amherst.
- Moulton, K. (2015). CPs: Copies and compositionality. *Linguistic Inquiry*, 46(2), 305–342. https://doi.org/10.1162/LING_a_00187
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.
- Radford, A. (2009). *Analysing English sentences: A minimalist approach*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511801618>
- Rizzi, L. (1997). The Fine Structure of the Left Periphery. In: Haegeman, L. (eds) *Elements of Grammar. Kluwer International Handbooks of Linguistics* pp. 281-337. Springer, Dordrecht. https://doi.org/10.1007/978-94-011-5420-8_7
- Rizzi, Luigi. 2004. Locality and Left Periphery. In A. Belletti, ed., *Structures and Beyond. The Cartography of Syntactic Structures*, vol.3, pp. 223-251. New York: Oxford University Press.
- Rosenbaum, P. S. (1967). *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.
- Shim, J. Y., & Ihsane, T. (2017). A New Outlook of Complementizers. *Languages*, 2(3), 17. <https://doi.org/10.3390/languages2030017>
- Swan, M. (2005). *Practical English Usage* (3rd ed.). Oxford: Oxford University Press.
- Tanigawa, S. (2018). Agreement, labeling and sentential subjects. *English Linguistics* 34(2), 302-330.
- Tegey, H., & Robson, B. (1996). *A reference grammar of Pashto*. Washington, DC: Center for Applied Linguistics.