

Why *sem* is (still) a complementizer and not a relative pronoun

Jóhannes Gísli Jónsson
University of Iceland

1 Introduction

This squib is about one grammatical word in Icelandic, the relative marker *sem*, as in (1a). As shown in (1b), relative clauses without a relative marker are ungrammatical in Icelandic:¹

- (1) a. Þetta er maður sem ég hef aldrei hitt.
this is man that I have never met
'This is a man that I have never met.'
b. *Þetta er maður ég hef aldrei hitt.
this is man I have never met.

Traditional grammars of Icelandic claim that *sem* is a relative pronoun, even though it does not inflect at all. Thráinsson (1980) disagrees and argues at length that *sem* is a relative complementizer, contrasting its behavior with relative pronouns that existed in older stages of Icelandic. Some of his arguments are familiar from the literature on relative markers in other languages, but some are more language-specific. Since the publication of the article, Thráinsson's view has become the received wisdom on relative *sem* in Icelandic as no scholar has attempted to rebut his arguments.

Since pronouns are quite different from complementizers, it would seem to be relatively easy to tease apart relative pronouns and relative complementizers. However, Kayne (2010) argues that the usual criteria for determining the lexical category of relative markers are questionable as they fail to divide them into two well-defined classes. He concludes that there is no principled distinction between relative pronouns and relative complementizers. This means, for example, that relative *that* in English, which is usually assumed to be a complementizer, can be argued to be a pronoun in basically the same sense as *who* and *which*.

Although Kayne (2010) is mainly concerned with English, he cites data from various other languages in support of his analysis. His arguments raise questions about the status of relative markers across languages, including Icelandic *sem*. This is an issue that I will address below where I will show that Thráinsson's (1980)

¹ The relative marker *er* is also used but only in highly formal language. As discussed by Thráinsson (1980), it has very similar properties to *sem*.

analysis of *sem* can be maintained and even reinforced to some extent. It turns out that *sem* can be argued to be different from the relative markers that Kayne (2010) builds his case on (see Sections 2 and 3) and it also displays some language-specific properties that lie outside the scope of Kayne’s analysis (see Sections 4 and 5). The overall picture that emerges from my discussion is that Icelandic makes a sharp distinction between relative pronouns and relative complementizers even though this may be blurred in many other languages.

2 Inflection

The examples in (2) show the contrast between *sem* and the relative pronoun *hver*, ‘who’, in Icelandic. The latter item inflects for case, gender and number whereas *sem* is invariant, irrespective of the gender and number of the head noun and the case associated with the relativized position:

- (2) a. Þetta er kona sem / hverja ég hef aldrei hitt.
 this is woman that / who.ACC.FEM.SG I have never met
 b. Hlauparinn sem / hver vann er frá Kenýa.
 runner.the that / who.NOM.MASC.SG won is from Kenya
 c. Börnin sem / hverjum hún hjálpaði voru þakklát.
 children.the that / who.DAT.NEUT.PL she helped were grateful

Although *hver* is hardly ever used as a relative pronoun in contemporary Icelandic, native speakers have very clear intuitions about its basic properties. For convenience, I will thus continue to construct Icelandic sentences with relative pronouns instead of referring to historical examples.

As a pronoun, *sem* would be highly unusual by the morphological standards of Icelandic. Thráinsson (1980) notes, however, that the simple third person reflexive only inflects for case (*sig*, *sér*, *sín*) and the *wh*-word *hvaða*, ‘which’, shows no morphological distinctions at all. He concludes from this that the lack of inflection is not a decisive argument against the pronominal status of *sem*.

I think the morphological argument against *sem* as a relative pronoun can be strengthened beyond these remarks. First, since the simple reflexive inflects for case, it can be maintained that *sem* cannot be a pronoun because all pronouns in Icelandic have morphological case. For Harbert (2007: 424), it is indeed case that is the crucial morphological property that characterizes relative pronouns in Germanic, as shown, for example, by English *who*. Second, the simple reflexive has the grammatical feature third person as it can only be referentially dependent on third person antecedents. By contrast, *sem* does not seem to have any person features as it can be used with third person antecedents, as in (2) above, as well as first and second person pronouns (in non-restrictive relatives). In this respect, *sem* also differs

from relative pronouns in Icelandic, which cannot occur in relative clauses headed by first and second person pronouns. This contrast is used by Thráinsson (1980) as an argument that *sem* is a complementizer.

The *wh*-word *hvaða* is not directly relevant to the issue at hand since it is a determiner and not a pronoun in that it always modifies a noun. Still, it is interesting in the present context as the only determiner in Icelandic that does not show any morphological distinctions for gender, number and case. It may nevertheless express some inflectional features through the suffix *-a*, which may be linked to the inflectional ending *-a* in the weak neuter singular which appears in all the four cases (cf. nouns like *hjárt-a* ‘heart’). (The root of *hvaða* is *hvað* ‘what’, so there is no doubt that *-a* is an independent morpheme.) This suffix may also be linked to the final *-a* that is found in various uninflected adjectives in Icelandic (cf. words like *gjaldþrota* ‘bankrupt’ and *örmagna* ‘dead tired’).

My point here is not to deny that uninflected relative pronouns may exist in some languages, as discussed by Kayne (2010), especially if they show some syntactic properties of pronouns. What I want to argue is that uninflected *sem* is not in that class since Icelandic is a morphologically rich language where every pronoun is at least marked for case. Note also that the lack of morphological case is only one of many arguments for the claim that *sem* is a complementizer. As we will see in the following sections, *sem* exhibits various properties of complementizers but no properties of pronouns.

3 Pied Piping

Possibly the strongest argument for Thráinsson’s analysis comes from the fact that *sem* can never pied-pipe a larger piece of structure, in clear contrast to relative pronouns. This is shown by the fact that *sem* cannot be the object of a preposition, as exemplified in (3b). Instead, the preposition must be stranded, as in (3a):

- (3) a. maður sem allir vilja tala við
man that everybody want talk to
‘a man that everybody wants to talk to’
b. *maður við sem allir vilja tala
man with that everybody want talk

Both preposition stranding and Pied Piping are possible with a relative pronoun. This is shown in (4).

- (4) a. maður hvern allir vilja tala við
man who.ACC.MASC.SG everybody want talk to

- b. maður við hvern allir vilja tala
 man with who.ACC.MASC.SG everybody want talk

The same argument can be made with possessives. As illustrated in (5), *sem* cannot pied-pipe a noun denoting a possessum, whether it precedes or follows the noun.

- (5) a. *kona sem börn ég þekki
 woman that children I know
 b. *kona börn sem ég þekki
 woman children that I know

By contrast, a possessive relative pronoun in Icelandic can pied-pipe the head noun, at least if the pronoun precedes the noun:

- (6) kona hvefarr börn ég þekki
 woman who.GEN.FEM.SG children I know
 ‘a woman whose children I know’

Kayne (2010) contends that Pied Piping is not a valid diagnostic for the grammatical status of relative markers. In support of this view, he points out that d-related relative pronouns cannot be objects of prepositions in Dutch, as shown in (7), even if they behave like pronouns in displaying number agreement in the neuter with the head of the relative, as in (8).

- (7) *de persoon aan die ik de brief heb gegeven
 the person to *die* I the letter have given
 (8) a. het boek dat ik nu lees
 the book *dat* I now read
 b. de boeken die ik nu lees
 the books *die* I now read

Kayne (2010) also observes that the relative pronoun *which* cannot pied-pipe a larger DP, as exemplified in (9):

- (9) *the book which’s first chapter is so well known

In my view, these data only show that the Pied Piping test must be applied with care, leaving room for the possibility that demonstrative-like pronouns may not pass it. In fact, since such pronouns in English lack a possessive form (cf. **that’s*, **this’s*, **these’s*, **those’s*), the ungrammaticality of examples like (9) is unsurprising. As for Icelandic *sem*, the important point is that it has no synchronic or diachronic relation to demonstratives or determiners that I am aware of. Thus, *sem* as a relative

pronoun would be expected to allow Pied Piping quite freely, contrary to fact and unlike Dutch *dat/die* and English *which*.

4 Other clauses with *sem*

Relative *sem* is believed to derive from the comparative complementizer *sem*, ‘as’, (Matthíasson 1959), a diachronic development that is also attested in German dialects (Brandner & Bräuning 2013). Thráinsson (1980) takes the origin of *sem* as yet another argument for its status as a complementizer because a change in the function of a complementizer is very natural whereas a change from a complementizer to a pronoun would be highly unlikely. In fact, relative pronouns are known to develop from *wh*-pronouns or demonstratives but not from complementizers (see, for example, Harbert 2007: 420–473 on Germanic).

Thráinsson 1980’s discussion of *sem* is mostly confined to relative clauses but it is also very instructive to examine other clauses where *sem* has a very similar function. These include concessive clauses, as in (10), which feature an initial *wh*-phrase as well as *sem*.²

- (10) Hvern sem þú velur mun ég styðja þig.
who.ACC.MASC.SG that you choose will I support you
‘Whoever you choose, I will support you.’

The most natural analysis of examples like these is to assume that the *wh*-phrase moves to a specifier position in the left periphery whereas *sem* is merged as a complementizer in the corresponding head position. An even stronger case for *sem* as a complementizer comes from exclamatives like (11a) and (11b) where *sem* interacts with verb second (Jónsson 2010).

- (11) a. Hvílíkur hálfviti sem hann getur verið!
what.NOM.MASC.SG idiot.NOM.MASC.SG that he can be
‘What an idiot he can be!’
b. Hvílíkur hálfviti getur hann verið!
what.NOM.MASC.SG idiot.NOM.MASC.SG can he be

The initial *wh*-phrase in (11a) is followed by *sem* but when *sem* is absent, as in (11b), the finite verb moves to second position. This verb raising is obligatory in the absence of *sem* and impossible when *sem* is present:

- (12) a. *Hvílíkur hálfviti hann getur verið!
what idiot he can be

² According to Van Riemsdijk 2006, clauses like these are not free relatives, despite appearances.

- b. *Hvílíkur hálfviti sem getur hann verið!
 what idiot that can he be

These data show that *sem* is in complementary distribution with V2 in exclamative clauses. This follows straightforwardly if *sem* is a complementizer, residing in a head position that the finite verb moves to when *sem* is missing. A pronoun occupying a specifier position in the left periphery would not be expected to interact with verb raising in this way.

5 Double complementizers

The final argument from Thráinsson 1980 to be discussed here concerns double complementizers. As illustrated in (13), the default complementizer *að*, ‘that’, can accompany various complementizers in colloquial Icelandic, including *sem*.³

- (13) a. Það eru margir sem (að) borða ekki skötu.
 there are many that that eat not skate
 ‘There are many that do not eat skate.’
 b. Ég fer ef (að) enginn annar vill fara.
 I go if that nobody else wants go
 ‘I will go if nobody else wants to go.’
 c. Þegar (að) hitt liðið skoraði breyttist allt.
 when that other team.the scored changed everything
 ‘When the other team scored, everything changed.’
 d. Hún spurði hvort (að) þetta væri löglegt.
 she asked if that this was legal
 ‘She asked if this was legal.’

By contrast, the complementizer *að* cannot occur with *wh*-phrases in embedded questions or relative clauses:

- (14) a. Hann veit hvenær (*að) María kemur.
 he knows when that Mary comes
 b. Þú ræður hvert (*að) við förum á morgun
 you decide where that we go tomorrow
 c. Það eru fáir hverjum (*að) hann treystir fullkomlega
 there are few who.DAT.MASC.PL that he trusts completely

³ I am not aware of any particular restrictions on *sem að* but this requires further investigation. Thus, it is possible in (10) and (11), supporting the claim that *sem* is a complementizer in these examples.

Since *sem* patterns with complementizers in this respect rather than *wh*-phrases or relative pronouns, Thráinsson (1980) takes this to show that *sem* is a complementizer and I agree with that conclusion. Presumably, the presence or absence of *að* with other complementizers boils down to the expression of the feature Fin, assuming a cartographic approach to the left periphery (Rizzi 1997). A single complementizer encodes both Force and Fin, thus simultaneously marking the type of the clause and its finiteness. In the presence of *að*, however, the higher complementizer expresses only Force but *að* expresses Fin.

References

- Brandner, Ellen & Iris Bräuning. 2013. Relative *wo* in Alemannic: Only a complementizer? *Linguistische Berichte* 234. 131–169.
- Harbert, Wayne. 2007. *The Germanic Languages*. Cambridge: Cambridge University Press.
- Jónsson, Jóhannes Gísli. 2010. Icelandic exclamatives and the structure of the CP layer. *Studia linguistica* 64(1). 37–54.
- Kayne, Richard. 2010. Why isn't this a complementizer? In Richard Kayne (ed.), *Comparison and Contrast*, 190–227. Oxford & New York: Oxford University Press.
- Matthíasson, Haraldur. 1959. *Setningaform og stíll [Sentence structure and style]*. Reykjavík: Bókaútgáfa Menningarsjóðs.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman et al. (eds.), *Elements of grammar: A handbook of generative syntax*, 281–337. Kluwer.
- Thráinsson, Höskuldur. 1980. Tilvísunarforöfn? [Relative pronouns?]. *Íslenskt mál og almenn málfræði* 2. 53–96.
- Van Riemsdijk, Henk. 2006. Free relatives. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, vol. II, 338–382. Oxford: Blackwell.