Minimal pronouns

Anders Holmberg

Newcastle University/University of Cambridge

On-Usa Phimsawat

Burapha University, Thailand

1. Introduction[1]

The following sentences exemplify the so called inclusive generic pronoun, overt in (1), covert in (2) and (3).

(1)   One shouldn't be afraid of making mistakes.   [English]   ´

(2)   Tämän koneen voi   hoitaa yhdellä kädellä   [Finnish]

        this machine   can.3SG operate with one hand

        'One can operate this machine with one hand.'

(3)   díawníi   ŋaan hǎa yâak   mâak thâa mây còb   trii   [Thai]

        nowadays job   seek difficult very if   NEG finish B.A.

         'To seek a job is difficult nowadays if one hasn't finished a B.A.'

It is called inclusive because the generic reference includes the speaker, the addressee, and other people. It is, thereby, the most general of pronouns, semantically.The question we will address is how this property is encoded in the feature make-up of the pronoun. There are basically two hypotheses. One is that it is the most richly specified pronoun, specified for first, second, and third person. The other is that it is the least specified one, therefore the least restricted one, allowing reference to the speaker, the hearer, and other people. We will

explore a version of the latter hypothesis, following Phimsawat (2011). The question is, what features does this minimally specified pronoun still have? A restriction that the inclusive generic pronoun has, at least in some languages, is that it can only include humans in its reference. We will show that this is true of some, but not all languages. Focusing on languages where the inclusive generic pronoun is a null category, we will demonstrate that there is a correlation between having subject agreement and having the reference of the inclusive generic subject pronoun restricted to humans. The task undertaken is to explain this correlation.

2.   Inclusive, quasi-inclusive and exclusive

The inclusive generic pronoun can be contrasted with the quasi-inclusive generic pronoun 'we', as in (4), and the exclusive generic pronoun 'they' as in (5).

(4)   We like smoked fish in Finland.

(5)   They died young in the Middle Ages.

Generic 'we' is called quasi-inclusive because it includes the speaker but not necessarily the addressee. (4) would typically be uttered by a Finn to a foreigner. It can be paraphrased as 'people in general in Finland, of which I am one'. Generic 'they' is exclusive in that it excludes the speaker and the hearer. The pronoun in (5) can be paraphrased as 'people in general in the Middle Ages'. The quasi-inclusive and exclusive generic pronouns both typically require the specification of a domain, either geographical or temporal, where the temporal domain typically denotes a historical period; see Holmberg and Phimsawat (2015).

In Thai, a radical pro-drop language, the quasi-inclusive pronoun has to be overt, in an out of the blue situation, as shown by (6).

(6)   **raw** kin   cee   nay dɨan tùlaakhɔm
         we   have vegetarian food in   month October
         'We have vegetarian food in October.'

With a null subject (6) would either be interpreted as inclusive generic ('One has vegetarian food...') or as having a referential 1ˢᵗ person subject ('I have vegetarian food...')..The quasi-inclusive pronoun can be null if it is bound or controlled by an overt one; see Holmberg and Phimsawat (2015).

(7) **raw** kin cee   nay dɨan tùlaakhɔm lǎŋ Ø thamboonsàjbàat

    we have veg. food in month October   after offer food to monk

  'We have vegetarian food in October after offering food to monks.

The exclusive pronoun can be overt or covert; see Holmberg and Phimsawat (2015) for more details.

(8) bon kɔ̀   níi   sùanyài *(khǎw)* plùuk chaa khǎay

    on island DEM mostly   they grow tea   sell

    'On this island they grow and sell tea.'

In this the exclusive and quasi-inclusive pronouns contrast with the inclusive pronoun, in Thai, as the inclusive pronoun can be null in out of the blue sentences, in fact must be, as there is no overt counterpart.

The present paper will focus on the inclusive generic pronoun. The quasi-inclusive and exclusive pronouns are mentioned here to show that they can be clearly distinguished empirically from the inclusive one.

3.   The inclusive generic pronoun in Thai has no phi-features

What features does an inclusive generic pronoun have? The meaning is 'people in general, including me and you'. It has, thereby, the most general reference of all pronouns. There are two hypotheses how to encode this property as phi-features: One is that it is the most richly specified pronoun, specified for first, second, and third person, however this is formally expressed; see Hoekstra (2010). The other is that it is the least specified one, therefore

allowing reference to the speaker, the hearer, and everyone else. A version of the latter hypothesis is proposed in Nevins (2007), where impersonal pronouns have an underspecified person feature;. see Fenger (2016) for discussion. We will assume another version of the latter hypothesis, according to which the inclusive generic pronoun has no phi-features in some languages, namely language without agreement, including Thai, while it has minimal phi-features in languages with agreement. Phimsawat (2011) argues, for Thai, that personal pronouns have the featural make-up (9) while the inclusive generic pronoun has (10);[2] see Déchaine and Wiltschko (2002), Holmberg (2005, 2010a,b).

(9)   [uD, [φ [N]]]
(10)   [uD, [[N]]

uD ('unvalued D') is a referential feature, which is valued either by a referential index, which may be assigned freely or under anaphoric binding, or else by quantificational binding. In generic pronouns, and generic expressions more generally, the feature is bound by a generic operator, an adverbial operator GEN_x (= 'It is generally true of x') in the C-domain (following Moltmann 2006). The phi-features include person, number, and in some languages, gender or class. We will discuss the properties of the feature/head N below. We will take this theory as a starting point. As we shall see, it cannot be the case universally that the generic inclusive pronoun is phi-featureless, because in some languages it triggers agreement.

As argued by Phimsawat, the absence of phi-feature specification explains why the inclusive generic pronoun is obligatorily null, in Thai: Having no phi-features means that there are no features to spell out, on the assumption that the uD feature and the categorial N-feature are, or at least can be, not associated with any phonological features.

This analysis of the inclusive generic pronoun is part of a theory, articulated in Phimsawat (2011), according to which arguments in Thai can be null if and only if (a) they

---

[2] In Phimsawat's (2011) notation the D-feature is R, for 'referential'.

have an antecedent which is sufficiently local, from which they can inherit a referential index, or (b) they have no phi-features but are bound by a generic operator.

An observation which can be explained immediately within this theory is that the quasi-inclusive pronoun cannot be null in Thai, in an out of the blue context. This follows since (a) the pronoun has the phi-feature value 1PL (excluding the addressee), and (b) being generic has no antecedent (see Holmberg and Phimsawat 2015). Since the value [1PL] cannot be deleted without irretrievable loss of information, it must be spelled out.

4.   Inclusive generic pronouns and reference to humans

We have said, and illustrated with examples, the claim that the inclusive generic pronoun includes the speaker, the addressee, and other people in its reference. What about inanimate things and non-human animals? Can they be included as well? Is it an integral property of the inclusive generic pronoun, or possibly generic pronouns more generally, that they only include humans in their reference, or is it just a consequence of the choice of predicates, so far? Predicates like 'be afraid of making mistakes', 'operate with one hand' and 'seek a job' select a human subject. It is clearly not the case that generic reference in general is restricted to humans: *Tigers are dangerous*, *Cars are expensive* are examples of non-human generic subjects.

If it turns out that inclusive generic pronouns are restricted to human reference, this should be encoded by some feature or features, following the logic of Phimsawat (2011). We could then not maintain the explanation that the inclusive generic pronoun is null because it has no restricting features.

We will start by considering what the inclusive generic pronouns look like in some other languages.

(11)

English:	*one*, *you*
Thamil:	*oruvan* [also 'one (person)'], Ø (with 3SG agreement)
Sinhala	*kenek* [also 'one (person)'], Ø
Swedish:	*man* [also 'man'], *du* 'you'

Turkish: *insan* [also 'human'], Ø (with 3SG agreement)

Japanese: *hito* [also 'human'], Ø

Italian   *si,*'REFL', *tu* 'you'

Finnish: Ø (with 3SG agreement), *sä* '*you*'

Brazilian Portuguese: Ø (with 3SG agreement), *você* 'you'

Basque   Ø (with detransitivized verb)

Thai: Ø

Chinese   Ø

Central Kurdish: *hamu kas* 'any person'

Vietnamese   *chung ta* ['you+me+others'], Ø

English is a representative of languages where the pronoun is a cognate of the numeral 'one'.

Other languages in this category include Thamil, where the commonest form of the overt generic inclusive pronoun is *oruvan*, which is the masculine form of the numeral 'one', which can also refer to women but not to non-persons. In Sinhala, too, the inclusive generic pronoun is *kenek* 'one (person)'. Swedish, Turkish, and Japanese represent languages where the overt form of the inclusive generic pronoun is a cognate of the noun 'human' or, as in Swedish, 'man'. Italian represents languages (including most Romance and Slavic languages) where a reflexive clitic *si* (or a cognate thereof) is used to express inclusive genericity.

(12)   a.   Si lavora   sempre troppo.   [Italian]

              SI work.3SG always too,much

              'One always works too much.'

        b.   W tymdomu   umiera   się spokojnie.   [Polish: Krzek 2012]

               in this   house die.3SG SIĘ peacefully

              'In this house one dies peacefully.'

It is debatable whether the reflexive pronoun itself is the generic pronoun, or whether it is a voice-related, detransitivizing category which serves to license a null generic pronoun; see Cinque (1988), d'Alessandro (2008), Krzek (2012, 2013). There are also languages where

6

the passive is systematically used to express inclusive generic meaning. An example is Standard Arabic; see Fassi Fehri (2009). Basque, which is included in (11), also represents languages where the generic reading is marked by a special, impersonal verb form.

Finnish, Brazilian Portuguese, Basque, Chinese, and Thai represent languages where the commonest form, which may be the only form, of the inclusive generic pronoun is null. Central Kurdish represents languages where there is no designated inclusive generic pronoun, but where a quantificational expression meaning something like 'everyone', 'anyone', or 'whoever' is used. Vietnamese represents a possibly less common form of the inclusive pronoun. *Ta* means 'you+me' and *chung* is a pronominal associative plural marker. This is, thus, quite explicitly an inclusive pronoun.

Many languages, but not all, have the 2SG pronoun as an alternative inclusive generic form, overt or null with 2SG agreement. Interesting though it is, we will put aside the 2SG generic pronoun in this paper; see Gruber (2013).

In some languages the generic pronoun can be overt or null. This is the case in Japanese, for example. This is not a matter of optionality: in some contexts the pronoun must be pronounced, in other contexts it can be null, even when not bound by another generic pronoun (Seiko Ayano, p.c.). It is at present unclear what determines the distribution of overt and covert inclusive generic pronouns. We leave this issue for future research.

The list in (11) indicates that humanness is common as a feature of the inclusive generic noun/pronoun, as several of the pronouns are etymologically derived from a noun meaning 'human' or 'man'. In Thamil, the masculine inflection restricts reference to humans (Thamil has 'semantic gender marking' where masculine and feminine can only refer to male and female persons, respectively). In Vietnamese, the associative plural of *ta* 'you+me' can only refer to persons. It is not necessarily the case that a generic pronoun which is derived from a noun meaning 'human' would be restricted to human reference, though, since it may have been grammaticalized as an even more generic pronoun, including also non-human referents. Whether this has actually happened is an empirical issue. We will return to it briefly in section 7.

To test whether the human restriction is endemic to inclusive genericity we need to employ a predicate which can be applied to a human as well as a non-human subject. Since

the inclusive generic pronoun always includes the speaker and the addressee (or it would not be inclusive), the predicate must be compatible with human reference. But for the purposes of this test, it must also be compatible with non-human reference.[3] One such predicate is 'grow'. Humans can grow, but so can animals and plants. It is conceivable that the word for growth in humans and plants might not be the same in all languages. However, in the languages we have looked at so far, the same verb can be applied to all living beings. The test sentence we will use is a version of (12):

(13)   One grows well, if one gets good care and a lot of nutrition.

The context would be a person proudly showing his garden to a visitor, offering the sentence as an explanation why the garden is so lush. The sentence is meant to be a generalisation over humans, animals, and plants. In English, (12) cannot be used in this way: the generic pronoun *one* can only refer to humans (which shows, incidentally, that the etymological link to a noun meaning 'human' is not a crucial factor).

In this paper we will, however, only consider inclusive generic constructions with a null subject. This is to test Phimsawat's (2011) hypothesis that inclusive generic pronouns are null because they have no phi-features. See Fenger (2012) for discussion of the features of overt generic pronouns.

Consider the following list of examples. The extension, humans only or humans and plants, is indicated. The sentences are meant to be uttered 'out of the blue', i.e. the subject should not be anaphoric.

---

[3] Hualde and Ortiz de Urbina (2003: 587-588) test whether the Basque generic construction with an impersonal verb form must have a subject with human reference. However, in their two test sentences they employ a verb meaning 'bloom' and a verb meaning 'bark'. The result is ungrammatical, from which they conclude that the construction must have a human subject. But the sentences could also be ungrammatical because (a) the construction cannot exclude reference to humans, to be inclusive generic, and (b) these predicates cannot be applied to a human subject. See the text for evidence that the impersonal verb form in Basque is restricted to human reference, though.

(14)　thâa dâayráb khwaamrák khwaamʔawcaysày　kôo　cá　too　rew　　　　[Thai]

　　　if　get　love　care　　　　　then FUT grow　fast

　　　'If one gets love and care, one will grow up faster.' [humans and plants]

(15)　rúguǒ néng huòdé gèng duō　de　yíngyǎng, nàme　huì　　zhǎng de

　　　if　can　get　more much DE　nutrition　then　be.likely grow　DE

　　　gèng　kuài

　　　more　fast　　　　　　　　　　　　　　　　　　[Mandarin Chinese]

　　　'If one gets more nutrition, one will grow faster.' [humans and plants]

(16)　yeongyangpwun -ul　　seopchwiha-myeon,　ppali　calaŋ-ta　　　　[Korean]

　　　nutrition　　　-ACC　take　　-if　　　　quickly grow.PRES DECL

　　　'If one gets more nutrition, one will grow faster.' [humans and plants]

(17)　vadi poshana　labuvuth　　　honthata hadai　　　　　　　　[Sinhala]

　　　more nutrition　get-PTCP-CON　well　　grow-PRS

　　　'If one gets more nutrition, one will grow faster.' [humans and plants]

(18)　Nếu　hấp- thụ　được　nhiều chất dinh -dưỡng,　　thì　sẽ　　[Vietnamese]

　　　 if　receive　obtain　many CLF nutrition　　　COND　FUT

　　　phát- triển　nhanh

　　　 grow　　　fast

　　　'If one gets much nutrition, one will grow fast.'　 [humans and plants]

(19)　Sitä　　kasvaa　　nopeammin jos saa　paljon ravintoa.　　　[Finnish]

　　　EXPL grow.3SG　quicker　　if　gets　much nutrition

　　　'One grows quicker if one gets much nutrition.'　 [humans only]

(20)   im meqablim harbe ahava ve maym az gdelim   maher.   [Hebrew]

        if receive.3PL much love and water then grow.3PL faster

        'If one gets much love and water, one will grow faster.'   [humans only]

(21)   Com boa alimentacão cresce   mais rápido.   [Brazilian Portuguese]

         with good nutrition   grow.3SG more quick

        'One grows faster with good nutrition.'   [humans only]

(22)   Behar bezala zainduz   gero, hemen ongi hazitzen da   [Basque]

        appropriately take.care.IMP after here   well grow.HAB is

        "If one is treated appropriately, one grows well here"   [humans only]

According to our informants, the Thai, Mandarin, Korean, Sinhala, and Vietnamese examples may well be said about plants as well as animals and (necessarily) humans. The Finnish and the Hebrew examples cannot include plants. The Brazilian Portuguese example is not acceptable for all speakers (some speakers want an overt pronoun here, which would be *você* 'you' to convey the inclusive reading), but for those who accept it, it can only refer to humans. [4] The Basque example also cannot include plants.

One salient property that distinguishes Mandarin, Korean, Sinhala, Vietnamese, and Thai from Finnish, Hebrew, Brazilian Portuguese and Basque is that the former set lacks subject-verb agreement.[5]

---

[4] Marcello Modesto (p.c.) has provided the following example from Brazilian Portuguese as a case where a null generic pronoun can refer to plants and animals as well as humans.
(i)   Se está vivo, um dia morre.
        if is alive   one day dies
        'Whoever/whatever is alive, will die one day.'
This means that Brazilian Portuguese and Finnish are not exactly alike in relevant respects, and suggests that the correlation between agreement and human reference is not universal. We will return to this case in section 7.

[5] Three other languages which have a null inclusive generic pronoun and agreement, and are reported to allow reference to humans only are Bengali (Wim van der Wurff, p.c.), Assamese (Hemanga Dutta, p.c.), and

Thamil provides some interesting evidence that agreement is, or at least can be, crucial.

(23)   a.   kooda satthu   kidaithaal,   nalla valarum

more nutrition get.PRTC.CON well grow.FUT.**3N**

'If they get more nutrition they will grow well' [plants, not humans]

b.   kooda satthu   kidaithaal,   nalla valaruvan

more   nutrition get-PTCP-CON   well   grow.FUT.**3SG.M**

'If one gets more nutrition, one will grow well.' [humans only]

c.   kooda satthu   kidaithaal,   nalla   valaramudium

more   nutrition get.PTCP.CON   well   grow.INF.**can**

'If one gets more nutrition, one will grow well.' [humans and plants]

The null subject in (23a) can only refer to plants and animals because the gender agreement on the verb is incompatible with human reference. The null subject in (23b) can only refer to humans, because the gender agreement on the verb is incompatible with non-human reference. In (23c), the head of the predicate is a modal auxiliary which does not show agreement. Now the null generic subject can refer to humans as well as animals and plants.

Why would agreement make a difference to generic reference in languages which do not show the kind of gender agreement on T that Thamil does, though?

The following is a possible hypothesis, which can, however, be rejected: In the languages without agreement the null subject in (14)-(18) is ambiguous between an inclusive generic pronoun referring to humans in general and an exclusive generic pronoun referring to plants (or non-humans) in general. This hypothesis can be rejected, at least in the case of Thai, on the grounds that there is no exclusive generic pronoun, null or overt, which would refer to plants/non-humans.

Icelandic (Halldór Sigurðsson, p.c.). For various reasons we don't have examples from these languages directly comparable with the nutrition examples in (15)-(20).

(24b) thîi kɔ̀   níi yùudiikindii
       at   island this live well
       'They live well on this island.'

This sentence cannot be taken to be an exclusive generic statement about plants or animals, only about people; see Holmberg and Phimsawat (2015). To refer to plants and/or animals, the subject would have to be overt.

 The following is another possible hypothesis, which can also be rejected. The subject in (14)-(18) is not a generic pronoun at all, but a multiply ambiguous referential pronoun: 'I', 'you', 'he', 'it', 'they', etc., covering all people, animals, and plants. This can be rejected because referential pronouns other than first person and in some circumstances second person cannot be null in out of the blue sentences; they need a topic antecedent in the immediate discourse context (Phimsawat 2011, Holmberg and Phimsawat 2015). A first person, and in certain cases, a second person subject, can be null in out of the blue sentences because, in informal terms, the speaker and the addressee provide contextual antecedents for the null subjects. In more formal terms, the null subject can be bound by a 'speaker feature' or 'addressee feature', a syntactic representation of the speaker and the hearer in the C-domain (Sigurðsson 2004, 2007; Holmberg and Phimsawat 2015).

5. Inclusive reference in languages with agreement

We assume a Chomskyan theory of agreement (Chomsky 2001). Subject-verb agreement is formally a set of unvalued phi-features of T, person, number, and in Hebrew also gender. These features need to be assigned a value in the course of the syntactic derivation. They are assigned a value by the subject DP, being the closest DP which is 'active', not having been assigned a Case by some independent means. The valued phi-features of T are spelled out as an inflection on the finite verb or auxiliary, in the languages under discussion here. If the unvalued phi-features are not assigned a value, the derivation will crash at PF, as they, and thereby the finite verb, cannot be spelled out.

 This means that there must be a null generic subject in the structure, which has inherently valued phi-features. The agreement in the Hebrew example shows that it has 3PL.M. In Finnish and Brazilian Portuguese it has 3SG.

Could the 3SG in Finnish and/or Brazilian Portuguese generic sentences be default agreement, though? Default agreement is well known from many languages, employed when, for some reason, the phi-features of T (in the case of subject agreement) cannot be valued by the subject DP. This could be because the subject DP is assigned Case independently, and is thereby deactivated, or because there is no subject DP. Default agreement is typically 3SG. This can be seen in the Finnish sentence (22):

(22)   Minun   pitää   ostaa uusi   auto.
          I.GEN   should.3SG buy   new.NOM car.NOM
          'I should buy a new car.'

Some predicates assign genitive case to the subject, in which case it cannot assign phi-feature values to T. In that case, the phi-features of T get the default value 3SG (Laitinen and Vilkuna 1993). This suggests that the 3SG agreement in construction with the inclusive null generic subject could be default agreement. The same could then be true of Brazilian Portuguese. However, as demonstrated in Holmberg (2010b), the default agreement analysis is not right for Finnish. The argument is based on the fact that default agreement and 'true' agreement, including 3SG agreement, have clearly different effects elsewhere in the clause: If the subject of a transitive verb does not trigger agreement the object will get nominative case, as in (22). If the subject does trigger agreement, which entails that the subject gets nominative case, the object will get accusative case, as in (23).

(23)   Minä   voin   ostaa   uuden   auton.
          I.NOM   can.1SG   buy   new.ACC car.ACC
          'I can buy a new car.'

As shown in (24), sentences with a null inclusive generic subject show the same variation as sentences with an overt subject, which is to say that the null subject triggers agreement just like an overt subject. In (24a) the predicate assigns genitive case to the (null) subject, hence it

does not trigger agreement, and the object has nominative case. The verb has the default 3SG form.

(24) a. Nyt pitää   ostaa uusi   auto.

now should.3SG buy   new.NOM car.NOM

'One should buy a new car now.'

b. Nyt voi   ostaa uuden   auton.

now can.3SG buy   new.ACC car.ACC

'One can buy a new car now.'

In (24b) the subject triggers agreement, which is 3SG because the generic subject is 3SG. In return, the subject gets nominative, and the object consequently gets accusative.

Under the present theory of agreement, the existence of subject agreement marking on the verb which can be shown not to be default agreement, is evidence that there is a subject, even though nothing is spelled out (in the case of Finnish there is no overt form of a 3SG inclusive generic subject), and shows what phi-features it has, while tests such as the nutrition sentence test, can be used to show what other restricting features it has. We take it that we have established that it has the feature [+Hum] (we will later provide a reason for taking it to be the value of a binary feature rather than a privative feature). There are other tests which can be employed to establish whether an understood, but covert subject is actually syntactically represented. Such tests have been applied to the Finnish inclusive generic pronoun, and have showed consistently that there is a syntactically represented subject (Hakulinen and Karttunen 1973, Laitinen 1995, 2006, Vainikka 1989, Vainikka and Levy 1999, Holmberg 2010b). This covert subject can bind anaphora, control a PRO subject in a purpose clause, and license agentive adverbials. See Holmberg (2010b) for examples, with details. There is consensus among the linguists who have worked on the inclusive generic construction in Finnish that it has a syntactically represented subject.

We can explain why there has to be a subject with phi-features in the languages with subject agreement. We have not explained why that subject must be restricted to human reference.

6. Explaining the relation between inclusive reference, phi-features and humanness

First, what we call Human in grammar would be more appropriately termed something like Conscious Being, to also include talking animals and extraterrestrials and other such imaginary entities which have crucial human properties. With this proviso, we will continue to use the label Human or [±Hum].

There are various ways to integrate the feature Human in the structure of pronouns. One is that this feature is a component of N, the nominal 'base' of nominal expressions, perhaps appropriately seen as the root of a pronoun, a minimal root. *He* and *she* would have the root feature Human, or [+Hum], non-human-referring pronouns like English *it* would have a [−Hum] root. We may want to make a distinction between pronouns that get their interpretation from an antecedent and pronouns that do not. In the former case the component N, the root component of the pronoun, may be taken to be a copy of the NP of the antecedent, deleted under identity with this antecedent (see Panagiotidis 2002, Elbourne 2008 for different versions of this idea). In the case of the generic pronoun, there is no antecedent. Therefore it needs a root of its own. The [+Hum] feature would provide this. The fact that the inclusive generic reading includes, by definition, the speaker and the addressee in the extension of the pronoun means that in the case of this pronoun, the feature [−Hum] is not an option.

But what is the connection with agreement? What about all the languages where the generic pronoun is so inclusive that it can include plants along with humans and animals? In this case the pronominal root would seem to be unspecified for humanness, [±Hum], allowing reference to entities of any kind. The generalisation that we want to express, though, suggested by our data, is that a pronoun cannot have phi-features without specification of the feature [±Hum].

The following is an alternative. First, the minimal root of a pronoun is, universally, [ENTITY]. Second, there are two ways that a pronoun can refer to everything and/or everybody: one is not to have any phi-features, hence no restriction. The other is to have minimal phi-features, just enough to satisfy the requirements of agreement, yet allowing reference to the speaker, the addressee, and a maximally general set of 'non-participants'. The feature [participant], widely assumed as part of pronominal systems, following Harley and Ritter (2002), distinguishes between speaker and addressee on the one hand , and

everyone/everything else on the other hand. In Harley and Ritter (2002) all the features are privative. Third person is when the feature [participant] is absent, i.e. 'third person is no person' (see Nevins 2008 for discussion). Such a system does not allow for a pronoun with phi-features which allow reference to the speaker, the addressee, and everyone/everything else. The system must include a feature which can be underspecified for person: [±participant] (see Nevins (2008) for other arguments that this device is needed). On its own, this feature will not exclude reference to non-human entities, and therefore must be supplemented by at least one more feature.

Assume that the phi-feature set of a pronoun has to include at least one specified feature. The pronominal phi-features are person, number, and class (Harley and Ritter 2002). The inclusive generic pronoun, although formally singular is not semantically singular. Arguably this rules out the use of a pronoun specified for singular number as an inclusive generic pronoun. Assume that the first division among the class features is between human and non-human, as seen in the many pronominal systems which make a distinction between human and nonhuman third person pronouns. The inclusive generic pronoun cannot be specified [−Hum], as it must allow inclusion of the speaker and addressee. But it can be specified as [+Hum]. The minimal feature make-up of a pronoun with phi-features which will allow inclusive, generic reference will therefore be [±Participant, +Human].[6]

This presupposes that the unvalued phi-features of T are, or at least can be, formally valued by this minimal phi-feature set, where the spell-out of the so valued T is the third person singular suffix on the finite verb (in most but not all of the relevant languages; in Hebrew it is plural). That is to say, the third person singular form that the finite verb has in Finnish, discussed in section 5, would be a form of default agreement after all, in that the subject valueing the features of T would not be specified for person or number, but only for class (the [+Hum] value), which, however, has no morphological effect in Finnish.[7]

---

[6] Hebrew is a language with a null inclusive pronoun which triggers plural agreement, an option which would appear to be consistent with the semantics of inclusivity. The idea that one specified feature is enough would then seem to predict that the inclusive pronoun in Hebrew could remain unspecified for [Hum]. The data we have indicates that this is a false prediction.

[7] According to the theory of null subjects in Holmberg (2010a,b), Roberts (2010b), based on the theory in Roberts (2010a), null subjects in languages with agreement are derived by copy deletion. The valued phi-features of T and the subject pronoun form a chain of two copies, where one, the subject, is deleted, provided its features are a subset of the phi-features of T. Since the subject, if it is third person, is valued for gender (i.e. class) in many languages, T must be valued for gender as well, for the subject to be deletable, even when this is not morphologically realised, as is the case in many languages. The notion that T has, or may have, an invisible class feature in languages with phi-features in T thus has independent motivation.

In languages without agreement, there is no reason why a generic pronoun would have to have any phi-features. All it needs is the root feature [ENTITY] and merged with it, the [uD]-feature. When the D-feature is bound by the generic operator this results in a reading which can be rendered as 'entities in general including the speaker and the addressee', the minimally specified DP giving the maximally inclusive reading.

7.   A prediction for generic PRO

The theory predicts that arbitrary/generic PRO as found in the subject position of non-finite clauses should not have its reference restricted to humans, in languages or constructions where non-finite clauses do not exhibit agreement, as is the case in English, for example. In the absence of agreement there is no compulsion for the subject to have any phi-features, and therefore no compulsion to be restricted at all, beyond ENTITY. If the non-finite clauses exhibit agreement, as they can do in some languages, we expect inclusive generic PRO in those clauses to be restricted to humans.

In order to test the prediction we need a predicate which can apply to humans and non-humans, for instance plants.

(28)   On a day like this it's important [PRO to get enough water].

Can this be said as an explanation for watering the houseplants repeatedly, or for starting up the expensive sprinkler system on a wheat field? Native English-speaking informants that we have consulted agree that it can be, although some report a moment of hesitation before the judgment. All agree that (28) contrasts clearly with the overt generic construction in (29), employing *one*, which can only have human reference.

(29)   On a day like this it's important [that one gets enough water]. (humans only)

(28) also contrasts with (30), although less clearly than in the case of (31).

(30)   On a day like this it's important [that you get enough water]. (humans only?)

Some informants report a difference between (28) when PRO refers to houseplants (marginally OK) and when it refers to a wheat field (not OK). We return to this point below.

However, predicates that take complements with generic PRO typically have an implicit, if not explicit, experiencer argument controlling PRO: 'It is important/good/necessary/etc. for X_i [PRO_i to ...]'. The issue whether PRO is restricted to human reference or not may then more accurately be the issue whether the implicit/null generic experiencer is restricted to human reference or not. Whether it does may be an interesting question but it has no immediately obvious consequences for the issue at hand, which is the relation between agreement, phi-features, and reference to humans.

A predicate taking infinitival complements in English which does not have an experiencer argument is *common*.

(31)   It's common [PRO to deteriorate with age].

The predicate of the embedded clause is selected to allow a human or non-human subject, which can even be inanimate. The question now is, can (31) be said, for example, by someone inspecting a leaking roof, as a generic statement including roof tiles along with people and any other entities subject to aging? The prediction made by the theory sketched in section 4 is that it can be. Informants consulted (a limited number) agree that it can be, although sometimes after a moment's hesitation. All informants agree that there is a contrast between (31) and (32a,b), with an overt generic subject, such that the latter would be restricted to human subjects, hence not felicitous in the leaking roof context.

(32)   a.   It's common that one deteriorates with age.
        b.   It's common that you deteriorate with age.

Some informants also, again, report a difference between (32a and b), such that (32a) categorically excludes anything but human reference, while (32b) can be used, perhaps with an element of jocularity, for at least some non-humans, such as house plants.

For the class of languages which have subject agreement and a null inclusive generic subject pronoun in finite clauses, we predict a difference between inclusive generic pro and PRO: The former should admit only human reference, the latter should be more permissive. We have tested this with Finnish. The context is a farmer saying either (33a) or (33b), as an explanation why he is starting up his expensive sprinkler system to water his wheat fields.

(33)a. Näin kuumassa ilmassa   on tärkeää   saada   tarpeeksi vettä.   [Finnish]

      so   hot.INE   weather.INE is important get.INF enough   water

      'In hot weather like this it's important to get enough water.'

  b. Näin kuumassa ilmassa   on tärkeää   että saa   tarpeeksi vettä.

      so   hot.INE   weather.INE is important that get.PRS.3SG   enough   water

      'In hot weather like this it's important to get enough water.'

The prediction is that there would be a difference between (33a), with the infinitival complement, and (33b), with the finite complement, such that (33a) would be appropriate in this context but (33b) would not. As it happens there was disagreement among the informants whether there was a clear difference between (33a,b). Some informants confirmed the prediction, but other informants accepted them both.[8]

However, as in the case of the English examples, we cannot tell whether this is a matter of features of the implicit experiencer argument or the null subject. The following is a test using the predicate *on yleistä* '(it) is common', which can be read without any implicit argument.

(34)   a.   On yleistä   rapistua   vuosien myötä.

             is   common deteriorate.INF years   by

             'It's common to deteriorate with age.'

      b.   On yleistä   että vuosien myötä rapistuu.

             is   common that years   by   deteriorate.3SG

             'It's common that one deteriorates with age.'

As above, the question is whether (34a,b) can include roofs as well as people in the generic reference. Among the limited number of informants consulted there was variation Some did not accept either of them as generic statements including roofs, while some accepted (34b) with that interpretation. This is not predicted by the theory articulated above. A more careful investigation will have to be left for future research.

---

[8] Thanks to Saara Huhmarniemi for her help with these tests.

In section 4, note 4, we mentioned a counterexample, provided by Marcello Modesto (p.c.) to the generalisation that a null inclusive 3SG generic pronoun in a language with agreement can only have human reference.

(35)   Se está vivo, um dia morre.   [Brazilian Portuguese]
       if is alive   one day dies
       'Whoever/whatever is alive, will die one day.'

This sentence can be said when talking about plants, animals, and or humans. The following is a similar Finnish example, provided by Pauli Brattico (p.c.):

(36)   Sitä   syntyy ja   kuolee aikanaan. Mikään/kukaan ei   elä loputtomasti. [Finnish]
       EXPL is.born and dies   some.day nothing/no-one NEG lives for.ever
       'One day you are born, one day you die. Nothing/no-one lives for ever.'

Brattico reports that he could use this sentence, for example, in conversation with a child when the family's cat or some other important living entity, even a house plant, is dying The extension of the relevant feature of the generic pronoun is thus not humans but something like 'humans and our associates'. It is perhaps significant that the predicates in (35) and (36) are 'be born', 'be alive' and 'die', predicates denoting defining properties of animate beings, including plants just as much as humans.

As mentioned, we will leave a more detailed investigation of generic pronouns and the human feature in non-finite constructions for future research.

8.   **Conclusions**

The starting point is the hypothesis, articulated in Phimsawat (2011), that the inclusive generic pronoun is the least specified nominal category, which therefore has the most general reference, including the speaker, the hearer, and everyone else. The observation is that there is cross-linguistic variation as to whether the pronoun is or is not restricted to humans. Focusing on languages which have a null inclusive generic pronoun in finite clauses, we have found that the null inclusive generic pronoun is restricted to human reference in some of them, but not all. The generalisation, based on data from primarily ten languages, five without agreement, four with subject-verb agreement, and one (Thamil) with or without agreement) is that the pronoun is restricted to human reference in the languages that have

subject-verb agreement in finite clauses. The explanation proposed is (a) in languages with subject agreement, i.e. unvalued phi-features in T, the inclusive generic pronoun has to have at least one specified phi-feature, to value the phi-features of T; (b) if the pronoun is to be inclusive, it cannot be specified for number, which entails that it must be specified for class; (c) if the pronoun is to be inclusive, i.e. include the speaker and the addressee, it must be specified [+Hum].

References

Ackema, Peter and Ad Neeleman. 2013. Person features and syncretism. *Natural Language and Linguistic Theory* 31, 901-950.

Biberauer, T., Holmberg, A., Roberts, I., & Sheehan, M. (2010). *Parametric syntax: Null subjects in minimalist theory*. Cambridge: Cambridge University Press.

Brody, M. (2013). Silent People: The pseudo-impersonal. In Johan Brandtler, Valéria Molnár, Christer Platzack (eds.), *Approaches to Hungarian: Volume 13: Papers from the 2011 Lund conference*. Johm Benjamins Publishing Company.

Cinque, G. (1988). On *si* constructions and the theory of *ARB* . *Linguistic Inquiry* 19: 521-81.

D'Alessandro, R. (2007). *Impersonal si constructions.* Berlin/New York: Mouton de Gruyter.

Elbourne, P. (2008). The interpretation of pronouns. *Language and Linguistics Compass*: 119–150.

Déchaine, R. M., and M. Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33 (3): 409–442.

Fassi Fehri, Abdelkader. 2009. Arabic silent pronouns, person and voice. *Brill's Journal of Afroasiatic Languages and Linguistics* 1: 1-38. Also in A. Fassi Fehri (2012) *Key features and parameters.* Amsterdam: John Benjamins, 257-290.

Fenger, Paula. 2015. How impersonal does one get ? A study of man-pronouns in Germanic. http://ling.auf.net/lingbuzz/002802

Gruber, B. (2013). *The spatiotemporal dimensions of person. A morphosyntactic account of indexical pronouns*. PhD dissertation. LOT, Utrecht University.

Hakulinen, A. and L. Karttunen. 1973. Missing persons: on generic sentences in Finnish. *Papers from the ninth regional meeting of thee Chicago Linguistic Society*: 157-71.

Harley, H. & Ritter, E. (2002). Person and number in pronouns: a feature-geometric analysis. *Language* 78:482-526.

Hoekstra, Jarich. 2010. On the impersonal pronoun *men* in modern West Frisian. *Journal of Comparative Germanic Linguistics* 13: 31–59.

Holmberg, A. (2010a . 'Null subject parameter'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 88–124.

Holmberg, A. (2010b . 'The null generic subject pronoun in Finnish: a case of incorporation in T'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 200–230.

Holmberg, Anders and On-Usa Phimsawat. 2015. Generic pronouns and phi-features: Evidence from Thai. *Newcastle and Northumbria Working Papers in Linguistics*, vol. 21. http://www.ncl.ac.uk/linguistics/research/workingpapers/Volume21.1.htm

Hualde, José Ignacio and Jon *Ortiz de Urbina*. A grammar of *Basque*. Berlin: Mouton de Gruyter.

Iwasaki, S. & Ingkaphirom, P. (2009). *A Reference Grammar of Thai*. Cambridge: Cambridge University Press. 16

Krzek, M. (2013a) Interpretation and voice in Polish SIĘ and –NO/–TO constructions. In I. K. Chahine (ed.), *Contemporary Studies in Slavic Linguistics*. Amsterdam: John Benjamins 185-198.

Krzek, M. (2013b). Generic subjects and Voice in Polish impersonal constructions. In C. S. Rhys, P. Iosad, and A. Henry (eds.), *Microvariation, Minority Languages, Minimalism and Meaning: Proceedings of t e Iris Networ in Form l Linguistics'*. Cambridge: Cambridge Scholars Publishing. 186-206.

Laitinen, L.2006. Zero person in Finnish. *Grammar from the human perspective.* In M.-L. Helasvuo and L. Campbell (eds.). Amsterdam: Benjamins, 210-31.

Moltmann, F. (2006). Generic one, arbitrary PRO, and the first person. *Natural Language Semantics* 14: 257–281.

Nevins, A. (2007). The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory* 25 (2), 273-313.

Panagiotidis, E. P. (2002). *Pronouns, clitics and empty nouns*. Amsterdam: John Benjamins.

Phimsawat, O. (2011). *The syntax of pro-drop in Thai.* Ph.D. dissertation, Newcastle University.

Roberts, I. (2010a) *Agreement and head movement: Clitics, incorporation, and defective goals.* Cambridge, MA: MIT Press.

Roberts, I. (2010b) A deletion analysis of null subjects. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 58–87.

Sigurðsson, H. A. (2004). The syntax of Person, Tense, and speech features. *Italian Journal of Linguistics* 16:219–251.

Sigurðsson, H. A. (2015). About pronouns. http://ling.auf.net/lingbuzz/001593.

Vainikka, Anne. 1989. *Deriving syntactic representations in Finnish*. Ph.D. dissertation: University of Massachusetts, Amherst.

Vainikka, Anne and Yonata Levy. 1999. Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17: 613-71.