

46. Indefinite Pronouns

Martin Haspelmath

1. Overview of the values

This chapter examines **indefinite pronouns** like 'somebody' and 'something' and distinguishes two major types. On the one hand, they may be closely related to the **interrogative pronouns** 'who' and 'what', as in Russian, where *kto-to* 'somebody' and *čto-to* 'something' are derived from *kto* 'who' and *čto* 'what' by means of the indefinite suffix *-to*. On the other hand, they may be closely related to the **generic nouns** 'person' and 'thing', as in Persian, where 'somebody' is expressed as *kæs-i* [person-INDEF] and 'something' is expressed as *čiz-i* [thing-INDEF]. In addition to the two major types, there are some languages with **special expressions** for 'somebody' and 'something', unrelated to anything else. An example is Abkhaz (Northwest Caucasian), where 'somebody' is *aj'ə* and 'something' is *ak'ə* (Hewitt 1979: 158). Theoretically one might expect further types, e.g. indefinites derived from demonstrative pronouns, but no other types have been found.

Languages may also show a **mixed type**, e.g. when 'somebody' behaves differently from 'something'. For instance, in Khmer the word for 'somebody' is *kè:*, which is also translated as 'person' and is unrelated to *nɛək na:* 'who', but the word for 'something' is *ʎvɔɣ(-mú:əɣ)*, which is closely related to *ʎvɔɣ* 'what' (Jacob 1968: 134–135, 294, 322, 325).

Finally, languages may lack nominal indefinite pronouns entirely and express the equivalent of 'someone' and 'something' by means of an **existential construction**.

Thus, Map 46 shows five different values (see Haspelmath 1997: 242–243 for an earlier version of this map).

@	1. Interrogative-based indefinites	194
@	2. Generic-noun-based indefinites	85
@	3. Special indefinites	22
@	4. Mixed indefinites	23
@	5. Existential construction	2
	total	326

In addition to the nominal indefinites 'somebody' and 'something', languages usually also have adverbial and adjectival indefinite expressions ('somewhere', 'somehow', 'for some

reason', 'of some kind') and occasionally we even find verbal indefinites ('do something'). Typically, these behave in the same way as the nominal indefinites. Thus, in Nivkh (Russian Far East), indefinites of all kinds are derived from interrogatives, and in French, indefinites of most kinds are derived from generic nouns.

(1) Nivkh (Panfilov 1962:253–7)

<i>sid</i>	'what'	<i>sid-lu</i>	'something'
<i>aŋ</i>	'who'	<i>aŋ-lu</i>	'somebody'
<i>řain</i>	'where'	<i>řain-lu</i>	'somewhere'
<i>řaɖ</i>	'which'	<i>řaɖ-lu</i>	'some'

(2) French

<i>un</i>	'one'	<i>quelqu'un</i>	'somebody'
<i>chose</i>	'thing'	<i>quelque chose</i>	'something'
<i>part</i>	'part'	<i>quelque part</i>	'somewhere'

Since many descriptions do not provide information on adverbial indefinites, this chapter is restricted to the nominal indefinites 'somebody' and 'something'.

2. The four types

2.1. Interrogative-based indefinites. In the majority of cases, interrogative-based indefinites are literally "based on" interrogative pronouns in the sense that they are overtly derived from them by the addition of a bound marker. A few examples are given in (3).

(3) a. Lak (Nakh–Daghestanian; Caucasus; Žirkov 1955: 74–75)

<i>cu-riw</i>	'somebody'	<i>cu</i>	'who'
---------------	------------	-----------	-------

b. Modern Greek

<i>ká-pjos</i>	'somebody'	<i>pjos</i>	'who'
----------------	------------	-------------	-------

c. Korku (Munda; India; Nagaraja 1999: 44)

<i>je-ka</i>	'somebody'	<i>je</i>	'who'
--------------	------------	-----------	-------

d. Southeastern Tepehuan (Uto–Aztecan; Willett 1991: 55)

<i>jaroi'</i>	'somebody'	<i>jaró</i>	'who'
---------------	------------	-------------	-------

This bound marker is typically related to expressions for 'be', 'want', 'perhaps', 'or', or 'also' (see Haspelmath 1997: ch. 6–7 for detailed discussion and diachronic explanations).

In other languages, indefinite and interrogative pronouns are identical. This possibility exists in some Slavic and Germanic languages (e.g. colloquial German *Ist da wer?* [is there who] 'Is there somebody there?'), and it is particularly widespread in Australian languages, e.g. Jaminjung (Jaminjungan; Northern Territory, Australia) *nanggayin* 'who; somebody' (Schultze-Berndt 2000: 45). It has been suggested that in these languages, indefinites and interrogatives are really identical and have the general meaning of "lack of information"; the specific interpretations as indefinite and interrogative arise from the larger construction or from the context (see Bhat 2000, 2004, as well as Mushin 1995 for Australian languages). If this is correct, these indefinites are not synchronically "interrogative-based", but it appears that at least diachronically, the interrogative use is always primary (Haspelmath 1997: 176).

2.2. Generic-noun-based indefinites. A few languages clearly have dedicated indefinite pronouns which are based on the generic nouns 'person' and 'thing'. Instead of the noun 'person', some languages make use of the numeral 'one', and not uncommonly, these are accompanied by an indefinite determiner of some kind. A case in point is Italian, which has *qualc-uno* 'somebody' and *qual-cosa* 'something' (from *qualche* 'some' plus *uno* 'one' or *cosa* 'thing'), and English is similar (*some-one/some-body*, *some-thing*), although the noun *body* does not have the meaning 'person' anymore when it is used by itself. For these languages, we have to say that there are dedicated indefinite pronouns, because not all properties of the indefinites are predictable from their parts and from general rules (for instance the fact that *something* is stressed on the first syllable, as opposed to *some thing*).

But for most languages with generic-noun-based indefinites, there is no good evidence available that these expressions are different from ordinary indefinite noun phrases. In fact, some descriptions explicitly deny that they are indefinite pronouns. For Kobon (Trans-New Guinea), Davies (1981: 156) says: "There are no specific indefinite pronouns. This is expressed by an indefinite noun phrase such as *bɪ ap* 'a man', 'someone', *nan ap* 'something'." In this chapter, such cases where we lack evidence for pronominal status have been lumped together with languages like Italian and English because it would have been very difficult

to draw a line between the two types. The evidence for pronominal status can be rather subtle. For instance, French *quelque chose* 'something' at first glance looks like a regular indefinite noun phrase consisting of the indefinite determiner *quelque* and the noun *chose* 'thing'. However, the notion 'something good' is expressed with a construction that is reserved for pronouns: *quelque chose de bon* (cf. *quoi de bon* 'what good'), whereas an ordinary noun phrase would be *quelque bonne chose* ('some good thing'). Unless a language has been described exceptionally thoroughly, such additional evidence is usually unavailable, but I suspect that it will eventually be found in more languages.

Thus, "pronominal status" was considered irrelevant for the purposes of this chapter, and indefinites were identified on purely semantic grounds: if a pair of expressions serves to render the notions 'somebody' and 'something', it is taken into account for this chapter, and if this pair contains the words 'person' (or 'one') and 'thing', then these indefinites count as generic-noun-based. Another possibility is that the indefinite determiner alone is used for 'somebody', as in Hausa (*wani/wata* 'some; someone', *wani àbù* 'something', *àbù* 'thing'; Newman 2000: 153–154).

Since many grammars of languages with generic-noun-based indefinites are simply silent about ways of expressing 'someone' or 'something', information often had to be extracted from dictionaries, example sentences and texts. Even so, the map is likely to show some reporting bias: the actual proportion of generic-noun-based languages is likely to be higher than shown here.

2.3. Special indefinites. Special indefinites are indefinite pronouns whose roots are synchronically unrelated to other roots in the language. In a number of Indo-European languages, their diachronic origin is clearly in interrogative-based indefinites, but they have lost their synchronic connection to interrogatives. Examples are Spanish *alguien* 'somebody', from Latin *ali-quem* [INDEF-who], and *algo* 'something', from *ali-quod* [INDEF-what]; Latin *ali-* was a clearly segmentable prefix, but Spanish *alguien* and *algo* are synchronically monomorphemic. Many of the Indo-European languages of South Asia are like Spanish in this regard.

2.4. Mixed indefinites. This type is not particularly interesting because it subsumes diverse subtypes: the 'somebody' word may be generic-noun-based while the 'something' word is

interrogative-based, as in Semelai (Mon-Khmer; Malay Peninsula): *smaʔ* 'person, someone', *mandemɔh* 'what; something' (Kruspe 1999: §6.2), or it may be the other way round, as in Maori: *wai raanei* 'somebody', *wai* 'who', *mea* 'thing, something' (Bauer 1993: 372). One of the two indefinites may be a special expression, e.g. Portuguese *alguem* 'somebody', not transparently related to *quem* 'who', vs. *alguma coisa* [some thing] 'something'. Or a language may have more than one set of indefinites, e.g. German: one set is interrogative-based (*irgend-wer* 'someone', *irgend-was* 'something') and another set is special (*jemand* 'somebody', not transparently related to *Mann* 'man', and *etwas* 'something', not transparently related to *was* 'what').

Finally, in a few languages the same morpheme may have all three senses: generic noun, indefinite, interrogative. Thus, in Urubú-Kaapor (Tupi-Guarani; northeastern Brazil), *awa* is 'person, someone, who', and *ma'e* is 'thing, something, what' (Kakumasu 1986: 331, 342, 380). These languages were also classified as mixed.

2.5. Existential construction. A language may completely lack nominal expressions for 'somebody' and 'something' and resort to an existential circumlocution instead. One of the two examples in my sample is Tagalog (Austronesian; Philippines), where 'somebody' and 'something' are expressed by an existential construction consisting of an existential particle plus a verb form:

- (4) Tagalog (Schachter and Otanes 1972: 276)
- | | | |
|--|-------------------------|-----------------|
| <i>May</i> | <i>d<um>ating</i> | <i>kahapon.</i> |
| exist | <ACTOR.VOICE>come.PFV | yesterday |
| 'Someone came yesterday (lit. There exists (one who) came).' | | |

3. Indefinite meanings

Most languages have different sets of indefinite pronouns for different meanings or contexts. The contrast in English between *some*-indefinites (*somebody*, *something*, *somewhere*, etc.) and *any*-indefinites (*anybody*, *anything*, *anywhere*, etc.) is fairly typical. The *some*-indefinites must be used in affirmative declarative sentences with realis modality (*Somebody left*/**Anybody left*), and the *any*-indefinites are used to convey a free-choice meaning (*You can take anything you like*, different

from *You can take something you like*) and can be used in questions and conditional clauses (*Has anybody left? If anything disappears,...*). Different languages divide up the space of indefinite meanings in different ways (Haspelmath 1997: ch. 3–4).

Important in the present context is the fact that different indefinite meanings do not behave alike with respect to the main parameter of variation discussed here. Indefinites that express free choice, and more generally indefinites that translate as *any*-indefinites in English, seem to have a greater tendency to be interrogative-based than indefinites that are used in affirmative declarative realis sentences (roughly corresponding to English *some*-indefinites). Thus, in Hausa, **realis indefinites** (as we may call them for short) are generic-noun-based, but free-choice indefinites are interrogative-based:

(5) Hausa (Newman 2000: 153–154, 622–623)

realis indefinites		free-choice indefinites	
<i>wani/wata</i>	'some, somebody'	<i>koo-waa</i>	'anybody'
		cf. <i>wàa</i>	'who'
<i>wani àbù</i>	'something'	<i>koo-mee</i>	'anything'
cf. <i>àbù</i>	'thing'	cf. <i>mèe</i>	'what'
<i>wani wurii</i>	'somewhere'	<i>koo-yàushè</i>	'anywhere'
cf. <i>wurii</i>	'place'	cf. <i>yàushè</i>	'where'

For the purposes of this chapter, only realis indefinites have been taken into account. In practice this means that indefinites translated as 'somebody/something' have been included, but indefinites translated as 'anybody/anything' have been disregarded.

4. Geographical distribution

The geographical patterns in the distribution of the two main types are striking. Practically all African languages have generic-noun-based indefinites, and so do almost all the languages of New Guinea and the small Pacific islands. In striking contrast to this, the languages of North America, Australia and Eurasia (apart from its western fringe) overwhelmingly show interrogative-based indefinites. The "special indefinites" type is scattered all over Eurasia, and the mixed type shows no clear pattern. This is one of the best examples of a feature whose

distribution seems to be completely independent of language families and exhibits areal patterning on a continental scale.