A corpus-based study of anaphora related to indefinite pronouns
Um estudo baseado em corpus sobre anáforas de pronomes indefinidos

Camila Ulrich

Sabrina Barreto

Resumo: *Os pronomes indefinidos da língua inglesa são palavras que substituem outras sem especificar quem estão substituindo. Uma dúvida que os estudantes de inglês têm é como fazer anáforas para referir-se a pronomes indefinidos, que podem referir-se tanto ao gênero masculino quanto ao feminino? Em inglês não há um pronome que faça referência a um gênero específico, nem como especificá-lo com o uso de artigos, como se faz em português. Nosso objetivo é compreender como são feitas anáforas de pronomes indefinidos. Para tanto, os seguintes pronomes serão analisados: 'someone', 'somebody', 'anyone', 'anybody'. Nossos dados para análise serão textos jornalísticos do Corpus of Contemporary American English (COCA). Observaremos as primeiras 50 ocorrências de cada pronome para ver quais apresentam anáforas, e quais palavras ou expressões são normalmente utilizadas para fazer anáforas de pronomes indefinidos.*

Palavras-chave: *anáfora, pronomes indefinidos, Inglês, Linguística de Corpus.*

Abstract: *Indefinite pronouns are words which replace words without specifying which word they replace. A question asked by English students is how to use anaphora to refer to indefinite pronouns, since they can refer to any gender? In English, there is not a pronoun that can refer to a specific gender nor how to define the gender through the use of articles as in Portuguese. Our objective is to understand which words or expressions are most often used to make anaphoric references to indefinite pronouns. The following indefinite pronouns were analyzed: 'someone', 'somebody', 'anyone', 'anybody'. Our data for analysis was extracted from newspaper texts in the Corpus of Contemporary American English (COCA). We observed the first 50 occurrences of each pronoun in the corpus to see which ones present anaphors and to see what words or expressions are most often used to make anaphoric references.*

Key words: *anaphora, indefinite pronouns, English, Corpus Linguistics.*

1   Introduction

Students of English as an additional language often have questions related to English grammar. Sometimes those questions are easily solved by grammar books. When you cannot find information in books, a good alternative for that is to search a corpus. That way we can see how the language is actually being used and solve our questions.

A question that we had as students of English is how do we make anaphora to refer to indefinite pronouns such as "someone", "anyone", "somebody", "anybody" and so on, since they can refer to any gender?

It is important to point out that English has a pronominal gap when compared to other languages, our mother tongue Portuguese, for example. The gap is that in English there is not a pronoun that can refer to any gender nor a way to define the gender through the use of articles such as it happens in Portuguese.

For example:

"In case anyone calls me, please tell＿＿＿＿＿I am busy."

How would we fill the gap? Our options seemed to be:

a) him

b) her

c) her or him

d) them

The best option would probably be letter c), since it incorporates both genders without being necessarily plural like letter d. However, we suspected that it was much more common to find examples that used the latter.

Thus, our objective is to understand what words or expressions are most often used to make anaphoric references to indefinite pronouns. For that, we chose the following indefinite pronouns to be analyzed:

Someone

Somebody

Anyone

Anybody

We would like to test whether the plural of the third person is most commonly used than the singular as an anaphor to refer to indefinite pronouns, or if there are other words or expressions used for that purpose.

Using the same example mentioned before, we believe that it is more common to use the third person plural form (them), based on occurrences of those pronouns that we have noticed in our everyday lives.

"In case anyone calls me, please tell them I am busy."

2　Review of Literature

As the object of this study is the anaphora of indefinite pronouns, we will explain some basic notions about this topic.

Indefinite pronouns are words which replace nouns without specifying which noun they replace. These pronouns can be divided in three groups: singular, plural or both. Some examples of singular indefinite pronouns are the forms "another", "anybody", "each", "everybody", "everyone", "everything", "little", "nobody", "no one", "nothing", "one", "other", "somebody", "someone", "something". For plural indefinite pronouns, we can mention "both", "many", "several". In the third group of indefinite pronouns – those which can be used with singular or plural-, there are the words "all", "any", "more", "most", "none", "some". Our objective is to identify what is the preferred way to make anaphoras of these pronouns in a corpus of newspaper texts.

Anaphora, according to the Stanford Encyclopedia of Phylosophy, is sometimes characterized as "the phenomenon whereby the interpretation of an occurrence of one expression depends on the interpretation of an occurrence of another or whereby an occurrence of an expression has its referent supplied by an occurrence of some other expression in the same or another sentence". We believe that the second definition is related to the study we will develop.

Below, there is an example of traditional anaphora:

(1)　The man went to the supermarket. *He* bought many apples.

Di Sciullo (2007) presents examples of cases where a pronoun is anaphorically related to a quantifier in a precedent sentence. "In (2), the existential quantifier can be the antecedent of a pronoun. In (3), the universal quantifier cannot be anaphorically related to the individual pronoun "he", whereas this is possible in (4), where an indefinite expression, a corporate number, can be reconstructed as a restrictor for the proposition including the pronoun(…)

(2)　Someone came to the meeting. *He* was expected to vote on the motion.
(3)　Everyone came to the meeting. *He* was expected to vote on the motion.
(4)　Everyone with a corporate number came to the meeting. *He* was expected to vote on the motion."(page 1, our numeration)

Likewise, other examples of anaphoric pronouns for indefinite sentences are found in Moltmann, (2006). He states that "unbound anaphoric pronouns, as in (5 and 6), relate to an indefinite noun phrase as linguistic antecedent, but are not in a position

to naturally act as a variable bound by the antecedent:

(5)   Someone broke in. *He* stole the silver.

(6)   If someone breaks in, *he* will steal the silver.

(7)   Someone might break in. *He* might steal the silver.

On any reasonable logical analysis of the sentences in (5) (respecting minimal syntactic conditions), *someone* cannot be analysed as an existential quantifier binding the variable that he may stand for" (MOLTMANN, 2006, page 199, our numeration).

3   Methodology

The observation of sentences that contain the pronouns above will be done with the use of the Corpus of Contemporary American English (COCA – http://corpus.byu.edu/coca/). This corpus contains more than 425 million words of text and is equally divided among 5 kinds of texts: spoken, fiction, popular magazines, newspapers and academic texts. Within this corpus, we will work only with newspaper texts.

The five examples below were found in this corpus and present the pronoun "someone":

(8)   Someone in Marshall's limo thought maybe Williams limo had been pulled over, and they decided not to stop to see what happened.

(9)   Someone would show up who hadn't run all weekend.

(10)   Someone's paying it, and that's going to impact their behavior.

(11)   Someone should be held accountable for their failure to protect the people of this city.

(12)   Someone who's willing to do it, just say,' I'm taking the reins' and throw out some dates and throw out some locations.

We will manually count how many times each definite pronoun appeared to relate anaphorically to an indefinite pronoun in the first 200 occurrences of each pronoun in the corpus. We will also compare the occurrences to see the most frequent word or expression used as anaphora related to indefinite pronouns in the newspaper texts of the corpus.

4   Analysis

The first pronoun analyzed was "someone". We typed the word into the search system of the corpus and selected the option of searching only newspaper texts. As a result, the system

displayed several parts of texts that had the word "someone". We then looked at each of the first 50 examples to see which ones had occurrences of anaphors referring to the indefinite pronoun.

The following anaphors were found in the corpus:

Who:

- As a result, he suggests, the sight of the young man's mother did not produce its usual emotional buzz, and his brain coped with this anomaly by rationalizing it as the presence of **someone who** looked like his mother but was in fact not her.
- For example, **someone** living in Spain **who** often visits family in France might purchase SIM cards for wireless services in both countries.
- It is hard to reconcile how **someone who** claims to love his country could give it such a bad name, steal from its coffers, abuse its institutions, pervert its constitution and undermine its quest for democracy.
- It's got to be **someone who** believes in being the leader of their peers.

That:

- One thing that I wanted to do is hire **someone that** could recruit the southern part of Georgia and also had some Kentucky ties.

Their:

- It's got to be **someone** who believes in being the leader of **their** peers.

Them

- No other professions will hire **someone** and expect **them** to do the same exact job as **someone** with 30-plus years of experience.

They

- **Someone** had turned off the water cutoff, and then **they** took a $5 copper fitting.

Her:

- " It's like Chinese water torture to lose **someone** and not know what happened to **her** day after day, year after year, "

His:

- Actually, calling this a bleat is unfair to sheep; it's more like **someone** with a head cold clearing **his** sinuses.

It:

- It's about time **<u>someone</u>** did this and we're just glad **it**'s not Apple.

Repetition of the pronoun "someone":
- Speaking out when **<u>someone</u>** is cheating or being bullied in the schoolyard - these are situations when **<u>someone</u>** decides not to be a bystander.

From the examples above we can see that there was a higher occurrence of the pronoun "who" in the first 50 appearances of the word "someone" in the corpus. It happens because "who" is working as a relative pronoun – in Portuguese, it is expressed by "que".

(13)   Alguém **que** morava com os pais e agora vive na Espanha deve sentir muita saudade.

The process of using a relative pronoun is traditional and very popular among speakers of any language. The difficulty people have is to choose a personal pronoun. We can see that, in the last eight sentences, the results show us that people vary a great deal when using personal pronouns to talk about "someone" as a subject.

The second pronoun analyzed was "anyone". The use of anaphora for this pronoun was much less frequent than for the previous pronoun ("someone"). From the first fifty occurrences of the pronoun in the corpus, only four sentences had anaphora:

It:
- Kenneth Branagh, who plays Laurence Olivier in the film, says that if **<u>anyone</u>** could get Monroe right, **it** would be Williams.

Who:
- Typically, someone might get a bonus when they deliver the goods, but I don't know **<u>anyone</u>** in the business **who** gets that kind of money **who** hasn't delivered something.

Your/You:
- Can **<u>anyone</u>** reading this letter say that during **your** teenage years **you** did not get into trouble?

Them/Their:
- For his part, Leonard is adamant that he doesn't tell **<u>anyone</u>** to leave sensible jobs or safe relationships. Instead, he just tells **them** to pursue what Tennyson refers to as " the gleam " -- the thing **their** heart goes out to.

As we can see in the sentences found until now, we have more occurrences of relative pronouns, though it is not our object of analysis. The third pronoun analyzed was "somebody". From the first fifty occurrences of the pronoun in the corpus, the following sentences presented anaphoric relation:

Their:
- "You can't replace looking **somebody** in eye and shaking their hand and getting to know each other personally"

Who:
- It's surprising at first for **somebody** who comes to the game and they don't know what Terrance is about
- As **somebody** who chases smallmouth in streams, they looked intriguing to me.
- Nichol said he has personally recommended hiring Fleming on occasion, " if I run into **somebody** who's got a problem.
- And I would like **somebody** who knows the issues, first of all, and who doesn't rely on staff for everything.
- They'll give people the option to move into Monroe, but for **somebody** who lives in the West - Denver, Seattle or Portland - Monroe, La., is not the most attractive option.
- It is difficult to find an unemployed person, or even **somebody** who knows someone who has lost their job.
- At 35 and with a family, he needs the instrument to hold up for a long while, which can gnaw at **somebody** who admits to being obsessed with time.

They
- Hopefully it's a positive development, hopefully there's **somebody** in charge now, and they're going to come out here and do right by all these Web 2.0 companies that are hanging by a thread.

Them
- I just give a picture of **somebody** else. Not to judge them in any way, but to say how they're feeling.
- I think it's more effective to show **somebody** else's point of view. You get a little more freedom to think like them and come up with a different point of view.

For the last pronoun analyzed – anybody -, we found the following sentences with anaphoric relation:

You

- We all try to defy age, but when **anybody** gets to be 35, 36 years old, you're going to slow down, I don't care who you are or how strong you are.

Who

- **Anybody** who does well in the private sector and goes into the public sector has to be prepared to do that and we are.
- The Nats will have to find out if they have **anybody**, or two or three somebodies, who has what it takes.
- I don't think you can find **anybody** who can give you a non-ideological answer to that
- I've called, but you can't ever get **anybody** who knows anything about your case.
- I didn't know **anybody** who didn't like him.
- **Anybody** who follows the L.P.G.A., and even those who don't regularly, the one thing they all agree on is that the personalities on the L.P.G.A. are second to none
- I haven't heard of **anybody** who is sad about it.

They

- I would never criticize **anybody** for making any choice that they believe is the best choice for their child.
- **Anybody** who follows the L.P.G.A., and even those who don't regularly, the one thing they all agree on is that the personalities on the L.P.G.A. are second to none.

Their

- I would never criticize **anybody** for making any choice that they believe is the best choice for their child.

5   Conclusions

Our objective in the study was to discover the most frequent pronouns used to make anaphora of indefinite pronoun and, according to our analysis of newspaper texts in COCA, we can see that the most frequent one is "who". As we explained before, this is a relative pronoun and it is largely used to make anaphoric relations.

When we analyzed the word "someone", we had the following *continuum*:

$$\boxed{\textbf{who} > \text{that} = \text{their/them/they} = \text{her} = \text{his} = \text{it} = \text{repetition}}$$

Considering only the personal pronouns, the most frequent one is "they" – the third person of the plural form. This result is in accordance with our expectation mentioned in the introduction of this article. This pronoun is also the most frequent form with "someone" and "somebody", as we can see in the board below, which shows us the number of occurrences with each form:

	WHO	YOU	**THEY**	**THEIR**	**THEM**
Somebody	**7**	0	**1**	**1**	**1**
Anybody	**7**	1	**2**	**1**	**0**

Thus, to sum up, to make anaphora of indefinite pronouns people will most commonly use either a relative pronoun (typically "who"), or some variation of the pronoun "they", or the pronoun "you".

After finishing our analysis, we can recognize that the quantity of data is an essential factor to make a linguistic study. In this way, Corpus Linguistics serves to give us many different alternatives of occurrences – which is very important to confirm or not our results. Therefore, a larger number of data would confirm more precisely our results.

References

DI SCIULLO, Anna Maria. *Pronominal Anaphora Processing.* World Academy of Science, Engineering and Technology. 2005.

CORPUS OF CONTEMPORARY AMERICAN ENGLISH – COCA. Available at: http://corpus.byu.edu/coca/. Accessed in Oct and Nov.

ENGLISH PLUS. Learn English with courses from Kaplan International Colleges. Available at: http://englishplus.com/grammar/00000027.htm. Accessed on: 15 Oct 2012.

MOLTMANN, Friederike. *Unbound anaphoric pronouns: e-type, dynamic and structured-propositions approaches.* Springer, 2006.

STANFORD ENCYCLOPEDIA OF PHILOSOPHY. Available at: http://plato.stanford.edu/ Accessed on: 16 Oct 2012.