**A construção do *corpus***

Definimos alguns procedimentos para a coleta de dados dos jornais. Inicialmente, criaremos um “dia artificial”, com base na proposta de uma semana artificial (BAUER, 2012). Assim, seguiremos a regra da representatividade (FONSECA JUNIOR, 2012, p.292), para constituir a amostra. As *home pages* dos jornais serão salvas com o auxílio do plugin *Capture Webpage Screenshot*, do navegador *Google Chrome.* Em publicação organizada por Palacios e Díaz Noci (2009), os pesquisadores defendem que é de suma importância o estabelecimento de critérios na seleção da amostra, também é preciso observar a representatividade da amostragem. Assim, na coleta de amostras considerou-se o horário em que a internet é mais acessada. Segundo relatório do Instituto Verificador de Circulação (IVC), publicado em 2012, durante os dias de semana e fins de semana, os horários de acesso a websites variam de acordo com os dispositivos utilizados (PCs, smartphones e tabletes):

“Os PCs concentram maior utilização no horário comercial. Os smartphones são usados de forma uniforme ao longo do dia e da noite, com picos nos horários de maior mobilidade: 8h, 13h e 19h. Já os tablets são utilizados de maneira uniforme no horário comercial, mas com intensificação durante a noite e no horário do rush matinal (8h). Nos fins de semana, os três dispositivos apresentam usos semelhantes: o dia começa mais tarde, termina mais cedo e todos possuem prime time às 20h” (Relatório IVC, 2012).

Com base no relatório e considerando a utilização de forma geral dos dispositivos no acesso a conteúdos digitais, constatamos que o horário em que os usuários mais acessam a internet é entre 8h e 20h. O esquema de escalonamento de horários de coleta pode ser visualizado na tabela a seguir:



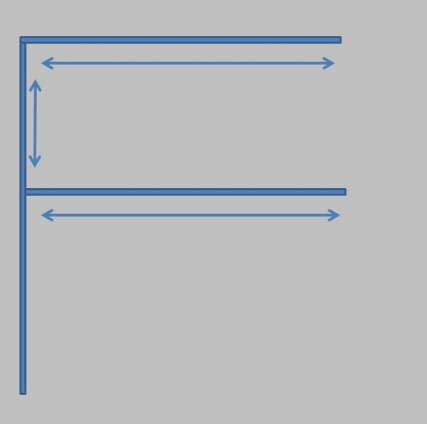
**Tabela 2:** Escalonamento de horários para coleta do *corpus*

**Fonte**: autoria prória

Pesquisas sobre o percurso da leitura do usuário na web para analisar o conteúdo também servem de aporte para a coleta dos dados. Segundo a pesquisa EyeTrack07 (2009), a direção dos leitores nas páginas começa no lado esquerdo superior da página inicial. O pesquisador Jakob Nielsen (2006) argumenta que o padrão de leitura dos usuários na *home page* é estabelecido em um movimento horizontal iniciado na parte superior esquerda, em direção ao lado direito da página. Só então, o usuário realiza o movimento vertical de leitura.

“Em alguns casos, os usuários irão ler através de uma terceira parte do conteúdo, tornando o padrão mais parecido com um E ao invés de F. Em outras ocasiões, a leitura será uma única vez, fazendo com que o padrão torne-se parecido com um L invertido (com a barra no topo). Entretanto, em geral os padrões de leitura assemelham-se a um F, embora haja variação na distância entre a barra superior e inferior” (NIELSEN, 2006, tradução nossa).[[1]](#footnote-1)

O padrão sintetizado por Nielsen (2006) é o mais adotado e se assemelha a um F, como pode ser observado na figura 5:



**Figura 5:** Padrão de leitura F

**Fonte:** FRANCO, 2009, p. 41. Adaptação nossa.

Com base no padrão F de leitura, realizaremos a coleta das notícias localizadas na *home page* dos portais Estado de Minas, Hoje em Dia e O Tempo. As amostragens seguirão o fluxo da esquerda para a direita na parte superior, local das notícias de maior destaque, dos três ciberjornais.

1. Texto original: “Sometimes users will read across a third part of the content, making the pattern look more like an E than an F. Other times they'll only read across once, making the pattern look like an inverted L (with the crossbar at the top). Generally, however, reading patterns roughly resemble an F, though the distance between the top and lower bar varies” (NIELSEN, 2006). Disponível em: <<http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/>> [↑](#footnote-ref-1)