

Considerações sobre a oralidade para estudos multimodais e análise de audiovisual televisivo

pg 25-37

Larissa Pinho Cavalcanti¹

Resumo

O presente artigo discute como a oralidade tem sido abordada em estudos multimodais sobre audiovisuais. Primeiramente, aborda a relação da fala e da escrita e o status da fala em produções televisivas, não compreendidas como manifestações naturais de oralidade, mas oralização de um roteiro já memorizado, com turnos simulados entre interlocutores. Então, discutimos como as transcrições multimodais organizam os dados verbais oralizados em produções audiovisuais em sua relação com os modos visuais e sonoros. Finalmente, utilizamos o modelo de análise de nosso trabalho doutoral para propor uma forma de transcrever dados de oralidade em audiovisuais, com exemplo de análise, também com corpus retirado de nossa tese. Com isso, esperamos contribuir para a elaboração de técnicas de transcrição multimodal linguisticamente orientadas.

Palavras-chave: audiovisual; oralidade; transcrição multimodal

ORALITY ISSUED IN MULTIMODAL STUDIES AND ANALYSIS OF TELEVISION AUDIOVISUAL

Abstract

The present paper focuses on how orality has been approached in multimodal studies on audiovisual genres. First, the relation between orality and writing is discussed both in general and given the context of television production, in which orality is not understood as manifestations of natural speech but the verbalization of a memorized script, enacted with simulated turn-taking and reactions. Then, the discussion turns to how multimodal transcription frameworks have been organizing the verbal content of audiovisual productions in their relations to both visual and aural modes. Finally, the model of analysis used in our doctoral thesis is shown as example of multimodal transcription for verbal data based on corpus also from our thesis. The elaboration of more comprehensive transcription and analysis techniques for language data is the expected outcome of the present work.

Keywords: audiovisual; orality; multimodal transcription.

Apresentação

Até meados da década de 1990, a investigação de sentidos gerados por semioses não verbais pode ser definida como exclusividade de áreas como design, artes plásticas, história da arte ou musicologia. Na linguística, mesmo os estudos que consideravam aspectos contextuais e informações não verbais tendiam a olhar semioses além do verbal apenas como “acessórios” cujos sentidos apoiavam ou se propunham de

1. Doutorado em Linguística pela UFPE, professora em Letras, na Universidade Federal de Pernambuco. E-mail laracvanti@gmail.com

modo independente ao mobilizado pelos conteúdos verbais. A evolução tecnológica e a diversificação das práticas comunicativas reconfiguraram o cenário das comunicações no século XXI a ponto de tornarem evidente a insuficiência de estudos linguisticamente orientados exclusivamente para o material verbal (escrito ou falado) para dar conta dos diferentes sentidos gerados por textos digitais ou audiovisuais.

Nesse contexto surgem os estudos de multimodalidade como uma expansão de horizontes para as práticas investigativas da linguística. Naturalmente, textos e interações sempre foram multimodais, o que muda com a perspectiva da multimodalidade é a emergência de um campo de pesquisas que se beneficia ao mesmo tempo das novas possibilidades de interação e das tecnologias digitais que expandem os métodos de observação, registro e análise de dados. A partir dessas noções gerais, os estudos em multimodalidade têm se caracterizado como um contínuo entre investigações que se dedicam ao estudo de modos específicos como a fala, a música ou o gesto e, mais recentemente, aquelas que olham para expansões semânticas na integração de texto e imagem, fala e gesto.

Com pesquisas tão diversas que se definem como multimodais, é preciso atentar que estudos de multimodalidade propriamente ditos respeitam os seguintes princípios (JEWITT, 2016):

1. A língua como elemento de um conjunto multimodal divide com outros modos o poder de gerar sentido.
2. A seleção e configuração de modos faz com que os significados produzidos sejam mais que “a soma das partes”.
3. Os modos são modelados socialmente e culturalmente, de maneira que seus recursos materiais carregam profundas orientações históricas e sociais das sociedades e suas culturas para cada signo e podem ser usados para trabalhos semióticos específicos.
4. O contexto e das demandas e interesses dos atores sociais influenciam a escolha dos modos para realização de processos específicos na geração de sentido em um evento comunicativo.

Para os propósitos de nossa tese de doutoramento, voltada para o estudo das representações sociais sobre as mulheres latinas em uma série de televisão, os estudos da multimodalidade possibilitaram tratar de modo integrado os discursos reproduzidos pela imagem (figurino, maquiagem, movimento corporal e expressões faciais), pelo som (das falas, do ruído ambiente e da trilha sonora), bem como pela linguagem técnica da própria filmagem (com enquadramento, iluminação). Nossa análise, então, investigou os sentidos desenvolvidos na colaboração dos modos como fruto de investimentos ideológicos subjacentes àquelas representações sociais.

Aqui, de modo específico, discutiremos como organizamos as análises das falas nas cenas da série para fins de transcrição multimodal. Para isso, realizaremos uma apreciação da relação fala e escrita, situando-a no contexto das produções televisivas; abordaremos brevemente duas propostas de transcrição multimodal para, daí, apresentarmos nossa proposta metodológica, com seus parâmetros; por fim, mostramos como a proposta de transcrição da fala ocorreu em nosso instrumento de análise multimodal, com um recorte de nossa tese. Vale salientar que ainda há necessidade de ferramentas de análise para estudos multimodais que contemplem não somente a descrição dos modos em seus propósitos específicos, mas que permitam discutir, à luz de questões sociais, os significados gerados a partir de suas várias relações.

A relação entre fala e a escrita: o que ocorre na televisão?

Os estudos linguísticos, muito em decorrência de sua vertente estruturalista, ignoraram uso da fala (ou o estudo da fala em uso) em favor de uma visão de sistema da língua que, quando se voltava para

textos orais, o fazia a partir de uma fala idealizada. De outro modo, estudos sobre uso real da fala não raro sofriam comparações com estudos do sistema escrito da língua. Nessa perspectiva, a primazia cronológica da fala era indeferida em favor de sua caracterização como fragmentária, pouco coesa e dependente de elementos paralinguísticos – uma consequência de estudos centrados na construção histórica da indispensabilidade da escrita para o bem social enquanto modalidade de inserção de linguagem humana (MARCUSCHI, 2007).

Todavia, desde a década de 1980 já se pode afirmar que as notórias divisões entre fala e escrita não mais sustentam os estudos linguísticos e o trabalho de desconstrução dessas polarizações já está bem consolidado. De fato, como Marcuschi e Dionísio (2007) argumentam, as práticas linguísticas se dão em textos orais ou escritos (certamente integrados a outras semiologias como gestos, olhares, imagens ou *layout*) e tanto fala como escrita apresentam suas variações, salientando que a escrita não é critério para categorizar línguas/culturas em mais ou menos primitivas:

não há razão alguma para desprestigiar a oralidade e supervalorizar a escrita. Também não há razão alguma para continuar defendendo uma divisão dicotômica entre fala e escrita nem se justifica o privilégio da escrita sobre a oralidade. Ambas têm um papel importante a cumprir e não competem. (MARCUSCHI; DIONÍSIO, 2007, p.15).

A partir da compreensão que fala e escrita, enquanto práticas discursivas, integram e estruturam as práticas sociais, é possível defender que ambas não competem por essas práticas, mas possuem “arena preferencial” advinda de sua história e papel na sociedade (MARCUSCHI, 1997). Consequentemente, tanto escrita quanto oralidade possuem graus de formalidade/informalidade e mantêm graus de implicitude e heterogeneidade e são ambas suscetíveis a investimentos ideológicos. Tais distinções são fundamentais para pesquisas que

assumem uma perspectiva multimodal, não somente por diminuir rigores hierárquicos entre modos, mas por salientar que os usos de cada materialidade semiótica dependem das práticas que requisitam recursos específicos. Por isso, a narração de um jogo de futebol usa padrões de volume e velocidade de voz diferentes de quem lê histórias para crianças.

As áreas preferenciais a que Marcuschi (2007) alude na compreensão das particularidades das manifestações orais e escritas da língua também não devem ser compreendidas como estanques. No exemplo do autor, a área jurídica, conhecida pela diversidade de documentos escritos (desde os códigos de lei à abertura de processos), também faz uso intenso e extenso de práticas orais, como as audiências. Se as práticas sociais recorrem a uma e outra manifestação da língua, a própria organização dos recursos linguísticos de oralidade e escrita devem ser vistos menos como polos estanques, mas como um contínuo cujas variações são multidimensionais, isto é, não lineares. Na organização desse contínuo é possível ver que “comparando uma carta pessoal em estilo descontraído com uma narrativa oral espontânea, haverá menos diferenças do que entre a narrativa oral e um texto acadêmico escrito” (MARCUSCHI, 2007, p.62).

No que diz respeito à fala, é preciso pontuar que a espontaneidade não é condição única e suficiente para todas as práticas sociais da oralidade, como exemplificam os telejornais (que são, na verdade, oralizações de textos escritos). Há marcas linguísticas do texto falado espontâneo que evidenciam seu planejamento passo a passo, sua construção articulada entre os participantes de uma interação, as relações de poder que os orientam. Essas marcas podem se tornar mais sutis à medida que a fala retome um texto já planejado.

Sobre a relação entre escrita e fala para a mídia, Lopez (2004) irá dizer que a oralidade da televisão é secundária, pois não é pura, mas acompanha ou deriva da escrita. Como exemplo dessa relação, é

possível resgatar o princípio da mídia televisiva, quando predominava o hibridismo entre teatro e rádio em programas levados ao ar sem considerar as possibilidades próprias daquela mídia para sua estrutura. Por outra perspectiva, Lopez (2004) argumenta que no telejornalismo, em particular, a mudança dos padrões de rádio para televisão envolveram principalmente a mudança no uso da fala para a televisão: se no começo das transmissões, os apresentadores não se preocupavam com o aparecimento dos textos escritos pelos quais se guiavam, com o passar dos anos, a televisão passou a buscar a naturalização da apresentação da matéria e, hoje, até ocorrem interações orais quase espontâneas durante a transmissão.

Além disso, é importante fazermos a ressalva que Lopez (2004) ignora a possibilidade de transmissões ao vivo e a própria diversidade de formatos da televisão contemporânea, que permite a espontaneidade da oralidade na interação entre interlocutores. Para esses contextos, não é possível se argumentar a favor de uma “oralidade secundária”, pois as interações, ainda que tenham temas pré-definidos e uma organização já conhecida pelos envolvidos e pelos telespectadores, demandam reações rápidas diante do que ocorre em campo ou no próprio estúdio, possibilitando, inclusive, o inesperado. O que não é uma verdade unânime para todos os formatos televisivos: a teledramaturgia, publicidade ou desenhos animados, por sua vez, raramente permite improviso.

De fato, para os formatos narrativos de entretenimento (tais como novelas, séries, minisséries) a relação da fala com o texto escrito é menos flexível: os atores devem memorizar as falas para reproduzi-las com determinado grau de precisão. Por isso, ao pretendermos analisar diálogos de narrativas televisivas, devemos ter sempre em mente que não se tratam de diálogos espontâneos, mas de interações criadas no campo da ficção e da escrita que buscam se aproximar de

uma interação oral natural. Na sessão seguinte, fazemos uma pequena descrição da oralidade em formatos televisivos narrativos contrastando com estudos sobre usos reais da oralidade.

O planejado da fala televisiva: uma simulação do uso real?

Para a caracterização das trocas dialogais da narrativa televisiva, é necessário entender um pouco sobre a organização e a recursividade da conversação natural. Para isso, muito contribuíram os estudos de análise da conversação, em particular a argumentação de Sacks et al (2003) a favor de uma organização própria das trocas faladas e da distinção de alguns elementos característicos da interação falada. São exemplos de tais elementos a troca de falantes (em turnos de fala) em ordem variável e de duração incerta, a possibilidade de concomitância de fala de diferentes interlocutores e transições entre turno sem intervalos ou com ligeiras sobreposições.

Nas interações televisivas, por sua vez, os turnos de fala são cuidadosamente organizados e, geralmente, as cenas possuem duas personagens centrais que dominam o enquadramento da filmagem e, portanto, o turno de fala. Além disso, não é comum que personagens quebrem o fluxo da interação com hesitações, repetições ou reformulações por problemas de produção ou de entendimento do que foi dito. Ademais, para os diálogos televisivos não é típico que os turnos se prolonguem para apenas um falante (mesmo os discursos proferidos por uma personagem são compostos de sentenças curtas e rápidas) nem que ocorram truncamentos bruscos entre os falantes.

No que diz respeito aos elementos multimodais da sequencialidade da conversação, Steinberg (1998 apud DIONÍSIO, 2007) descreve os recursos não-verbais normalmente empregados pelos falantes de dada língua numa conversa,

notadamente: os sons que não fazem parte do sistema sonoro da língua; os movimentos do corpo como gestos, postura, expressão facial e olhar; a distância mantida entre os interlocutores e como eles se tocam ou não durante a interação, bem como o uso do próprio silêncio. Para a televisão, o silêncio ocorre para chamar atenção para alguma ação desenvolvida em cena e raramente como elemento constitutivo das interações – quando ocorre, é de modo breve para pontuar desacordo ou raiva com o interlocutor. De outro modo, a realização de sons não linguísticos também não é parte da oralização esperada em uma série de televisão, cujos sons são geralmente trilha sonora, ruído ambiente e a fala das personagens.

Os pontos de convergência entre os diálogos televisivos e as conversas mais espontâneas, por sua vez, são responsáveis pela verossimilhança das interações roteirizadas e do próprio universo diegético. Por isso, é importante que narrativas de época ou que apresentem situações de assimetria (de poder, de geração, de origem social) usem de modo distinto aspectos linguísticos como léxico e sintaxe, e simulem a troca de turnos e as atitudes dos interlocutores de modo realístico para os telespectadores.

Nesse ponto, *quem* é representado e *como* se dá tal representação na televisão permite questionamentos quanto ao investimento ideológico da produtora/emissora, afinal, as representações televisivas sobre membros de minorias sociais podem dar continuidade a estereótipos através dos diversos discursos constituídos audiovisuais/verbais. Uma vez que não é somente a concepção da narrativa, nem os diálogos em sua natureza verbal que reproduzem as representações, outrossim, uma conjunção de discursos multimodalmente articulados, damos continuidade a nosso trabalho com uma breve discussão sobre propostas de transcrição multimodal.

Transcrição multimodal e a fala: dos modelos existentes à nossa proposta

O corpus de nossa pesquisa é multimodal, articulando principalmente imagem, som e fala, isto é, um corpus multidimensional, com grupos de recursos semióticos complexos coordenados entre si para geração de sentidos. Para transcrição e análise dos dados, necessitamos, então, de ferramentas que permitissem analisar o material audiovisual sem abrir mão das características específicas do formato da série (blocos de cenas e fragmentação episódica da narrativa; múltiplos núcleos narrativos; técnicas de edição) o que se mostrou ser fora do escopo dos modelos de descrição e análise. Duas propostas que se aproximaram de nossos objetivos e demandas foram encontradas nos trabalhos de Baldry e Thibault (2001) e Burn e Parker (2003).

A primeira proposta de transcrição e análise multimodal, desenvolvida por Baldry e Thibault (2001, 2006), é chamada de transcrição multimodal e surge como um modelo teórico e uma abordagem analítica de cunho funcional para textos inseridos ou projetados em alguma superfície tecnológica (como a tela do computador, da televisão ou, ainda, a página impressa). Baldry e Thibault (2001) caracterizam os textos audiovisuais, como constelações variáveis e diferentes de som, imagem, gesto, texto e língua, cuja análise vai além do estudo de agrupamentos de modos.

Para os autores, os textos em sua natureza multimodal integram seleções de diferentes recursos semióticos em sua organização, isto é, os modos não ocorrem em isolamento em um texto, mas são combinados e integrados para formar um complexo que não pode ser explicado como a “soma dos significados de cada uma de suas partes”. Com isso, é possível entender o efeito que a interação de recursos de pequena escala pode ter sobre os altos níveis escalares de observação e interpretação de sentidos (BALDRY; THIBAUT, 2006), implicando em uma economia cognitiva.

De modo geral, a transcrição multimodal é organizada em (1) tempo (2) espaço visual (3) imagem (4) ação cinésica (5) trilha sonora e (6) interpretação multifuncional. O tempo, diz respeito à sequência temporal do material analisado, isto é, a marcação em segundos da gravação, seguindo o disposto por um dado software de reprodução de vídeos. O espaço visual e a imagem dizem respeito ao que é visto pelo telespectador: a imagem propriamente dita e sua descrição, respectivamente. A chamada ação cinésica diz respeito aos movimentos do corpo, agrupando diferentes arranjos espaços-temporais dos participantes do evento discursivo. Por fim, a interpretação multifuncional é a compreensão de todas as multifunções de todos os atos semióticos – a qual os autores lembram não ser exaustivamente detalhada, apenas um conjunto de observações sobre a saliência² de um modo semiótico empregada em uma dada fase ou subfase.

A transcrição multimodal foi uma das primeiras propostas dentro dos estudos multimodais a se voltar para gêneros audiovisuais. Porém, o excesso de notações descritivas (sem dúvida um dos aspectos mais elaborados da proposta) tornaria a leitura das análises cansativas por exigir constante memorização ou consulta aos significados das notações. De outro modo, a própria interpretação dos usos dos recursos a partir da transcrição é pouco destacada pelos autores que oferecem exemplos incipientes a partir de uma perspectiva funcionalista da linguagem.

A segunda proposta de transcrição e análise de Burn e Parker (2003; 2010) toma como ponto de partida a compreensão da imagem em movimento

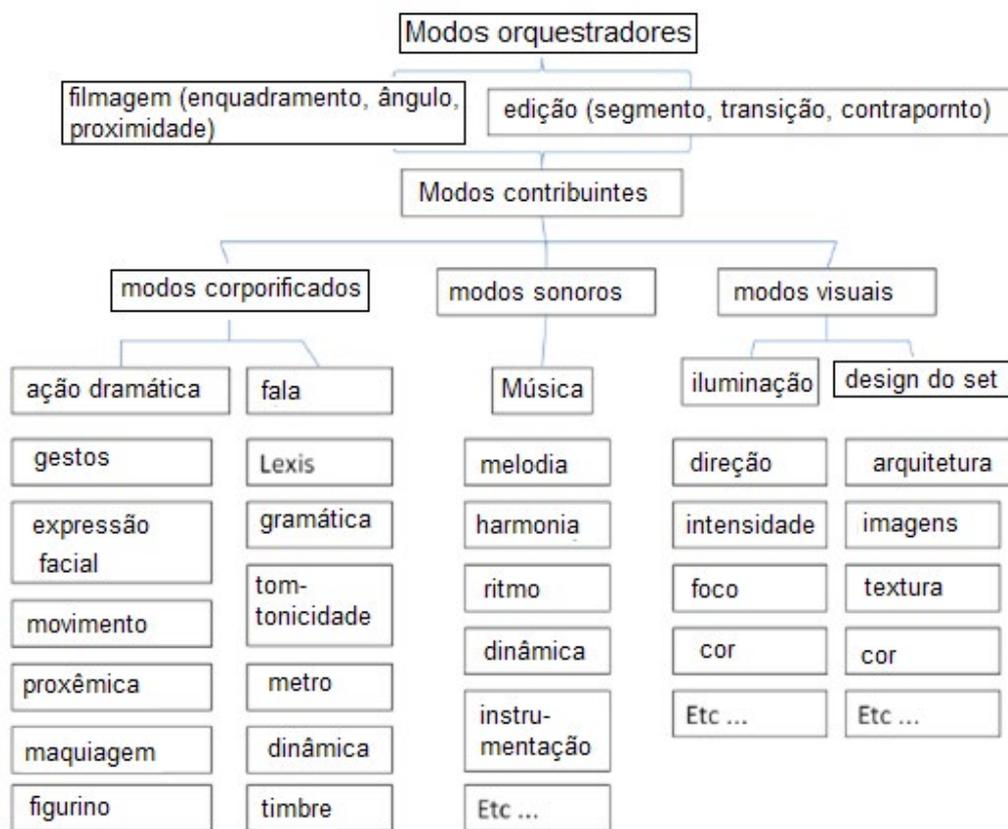
2 Para discussões mais aprofundadas sobre o conceito de saliência multimodal, ver Norris (2004).

como um conjunto combinatório e integrador de modos, que não deve ser limitada pela nomenclatura ‘fílmico’ ou ‘cinematográfico’. Diante dessa realidade, os autores adotam o nome kineicônico para abranger todos os elementos que se unem à imagem em movimento e são manipulados por práticas de filmagem e edição.

É a partir da reorientação promovida pelas técnicas de filmagem e edição que Burn e Parker (2003) e Burn (2010, 2013) irão distinguir, então, os modos orquestradores (os sistemas estruturadores de espaço e tempo) e contribuintes (movimento, luz, figurino, objetos, cenários, etc.) de um material kineicônico. A hierarquia apresentada por Burn (2013, p. 8), e adaptada na Figura 1, ajuda a visualizar como os modos contribuintes se organizam em função dos recursos de que dispõem e como elemento dos modos orquestradores. Para os autores, filmar e editar estruturam os materiais e sistemas de significação das artes visuais em relações de complementação, reiteração, antecipação ou contradição e, por isso, devem ser pensados como recursos que modalizam os sentidos dos demais modos.

Burn e Parker (2003) observam que os princípios que regem a descrição kineicônica irão operar de modo diferente em cada modo e através de diferentes modos. Posteriormente Burn (2013) irá salientar que o diferencial da análise multimodal é seu olhar para como os modos trabalham juntos, ou seja, não se trata de decompor modos semióticos maiores em elementos cada vez menores, mas de entender os recursos que conjuntamente produzem determinados efeitos de sentido.

Figura 1: Hierarquia de modos orquestradores e contribuintes



Fonte: adaptado de Burn (2013)

Apesar das distinções feitas pela Teoria Kineicônica, o modelo não é livre de falhas – uma das quais era a falta de espaço para análise do modo verbal oralizado no material audiovisual. Foi necessário, então, buscar um modelo que abrisse espaço para análise específica do componente verbal. A partir do modelo de Pereira (2014), foi possível, então organizar uma estrutura de descrição e análise que, sem abrir mão dos dados multimodais aplicável a audiovisuais de modo geral, permitisse a análise detalhada da fala. O Quadro 1 abaixo representa a organização final dessa proposta. Na primeira linha, organizamos os elementos de contextualização e identidade e visual das personagens ao passo que nas colunas distribuímos, em função do tempo (modo orquestrador), os modos de fala e ação dramática (gestos, expressões, deslocamento pelo cenário), com a imagem “resultante” do trabalho coordenado

desses modos. Nas linhas subsequentes, inserimos o modo orquestrador da filmagem e o contribuinte de trilha sonora, de natureza mais técnica.

Quadro 1: Quadro da estrutura de transcrição e análise para audiovisuais

Cena XX - Episódio XX			
Cenário			
Personagens			
Figurino			
TEMPO	FALA	AÇÃO DRAMÁTICA	IMAGEM
Modo orquestrador da filmagem			
Trilha sonora			

Fonte: Cavalcanti (2017) com base em Pereira (2014) e Burn (2013)

No que diz respeito à transcrição da fala, foco deste trabalho, optou-se por se utilizar apenas as normas de transcrição do NURC³, recorrentes em trabalhos linguísticos brasileiros, pois nosso material de análise era uma série norte-americana, ou seja, as normas de transcrição deveriam contemplar especificidades da expressividade oral da língua inglesa. Nos preocupamos, em particular, com aspectos de natureza prosódica como entonação e ênfase (*sentence* ou *word stress*), volume da fala, bem como com o alongamento de sílabas e as hesitações, uma vez que podem orientar a interpretação do telespectador. Para isso, baseados em Norris (2001), estabelecemos que:

Quadro 2: Convenções de transcrição

1. as hesitações	(...) quanto mais longa a pausa, mais pontos
2. alongamento de sílabas	repetição da letra associada ao som vocálico; o número de letras sugere a duração do prolongamento
3. a entonação crescente ou decrescente	uma linha reta diagonal próximo à fala
3. Sentence stress	linhas curvas próximas à fala cujo pico representa o ponto de maior ênfase
4. Word stress	Palavra sublinhada
5. Volume ascendente ou decrescente	Tamanho crescente ou decrescente da fonte

Fonte: Cavalcanti (2017) com base em Norris (2011)

Além disso, questões de intencionalidade do falante eram consideradas na própria análise, levando em conta o local da interação (ambiente doméstico, profissional ou externo, como mercados), a intenção dos interlocutores na interação (repassar informação por conversas ou fofocas, argumentar em discussão ou debate, etc.) e a relação entre as

personagens (familiares, profissionais, amorosas). A seguir, mostramos como essa proposta de transcrição e análise permitiu o estudo das falas de uma série de televisão.

Um exemplo de transcrição e descrição em séries: as falas em *Devious Maids*

A série *Devious Maids*, lançada em 2013 pela emissora Lifetime, é centrada na vida de cinco mulheres latinas (Marisol Suarez, Rosie Falta, Carmen Luna, Zoila e Valentiza Dias) que trabalham como empregadas domésticas para os ricos e famosos em Beverly Hills. Por questões de espaço, o exemplo deste trabalho se concentra em apenas um dos núcleos narrativos, o de Zoila e Valentina. Zoila é descendente de imigrantes há algumas gerações nos Estados Unidos e herdou de sua mãe a função de empregada da residência de Genevieve Delatour, uma socialite famosa por seus diversos casamentos. Por sua vez, Zoila leva a filha, Valentina, para trabalhar nessa mesma residência para que ambas possam pagar as mensalidades da faculdade de moda que a jovem deseja cursar.

A escolha pelo núcleo narrativo Zoila/Valentina se baseia na brevidade da cena analisada (o clímax de um conflito entre ambas) bem como no uso que ambas as interlocutoras fazem de diferentes recursos da oralidade para enfatizar seus pontos de vista. Antes da cena transcrita ocorrer, o conflito entre mãe e filha surge em virtude da proposta de Genevieve Delatour de usar um óvulo de Valentina para a geração de um herdeiro com seu mais novo enlace romântico. Zoila recusa a proposta, todavia, Valentina defende a possibilidade de pagar a faculdade através da venda de seu óvulo. Na cena analisada, o conflito atinge seu clímax.

Na Figura 2, há o primeiro excerto de transcrição da fala, respeitando original em inglês, com a tradução em rodapé. A coluna de imagem foi suprimida por questões de resolução e direitos de

imagem. Nesse primeiro momento da cena, Valentina se aproxima da mãe com a intenção de comunicar que conseguiu um meio de financiar sua educação, levando Zoila a uma conclusão equivocada. As ações corporais de Valentina demonstram insegurança e uma atitude defensiva, o que também é expressa em sua fala lenta e marcada por pausas. Zoila, por sua vez, é assertiva (note que usamos itálico para marcar palavras do espanhol) e não desvia a atenção de sua tarefa. Assim, vemos que os sentidos gerados durante a transcrição da fala ganham ênfase e completude com a descrição da ação corporal das personagens.

Figura 2: Excerto de análise 1

TEMPO	FALA	IMAGEM
00:21:00,140 - 00:21:02,700	So... I have good <u>news</u> .	Valentina entra na cozinha abanando os braços e batendo uma palma. Ela chega ao balcão e apoia as mãos sobre o mesmo. Zoila está removendo vegetais das sacolas de compras.
00:21:03,190 - 00:21:05,610	Well, don't just keep it to yourself. <i>A ver</i> .	Zoila está removendo vegetais das sacolas de compras.
00:21:05,980 - 00:21:08,220	I'm gonna <i>start fashion school</i> next fall.	Valentina torce as mãos em frente ao corpo.
00:21:09,540 - 00:21:11,910	They <u>changed their mind</u> about the <u>scholarship!</u>	Zoila abre os braços com as mãos viradas pra cima enquanto se locomove até Valentina para abraçá-la.
00:21:12,900 - 00:21:17,160	Not exactly. <u>I found a way</u> to pay the tuition.	Zoila larga a filha rapidamente para encará-la. Valentina inclina a cabeça para a direita, mexendo os cabelos e assumindo a autoria da mudança de rumo de sua vida.
00:21:18,210 - 00:21:25,720	- <u>How did you do that?</u> - Mr. Pettigrove came to me. He <u>told me</u> about Mrs. Delatour and what she asked you.	Zoila se afasta da filha com a cabeça levemente inclinada para trás reforçando seu afastamento do que é dito por Valentina.
00:21:25,755 - 00:21:26,685	Oh, my God.	Zoila dá um passo à esquerda.

Fonte: autoria própria com base em Cavalcanti (2017) ⁴

Animada com a possibilidade de a filha entrar na faculdade, a voz de Zoila aumenta em volume e também em *sentence stress*, com as palavras finais sendo mais enfatizadas que as palavras iniciais, e sua empolgação é representada com o gesto de largar a feira e abraçar a filha. A negativa de Valentina faz com que a mãe a afaste e Valentina usa o distanciamento para assumir uma postura mais assertiva. Com isso, há espaço para a ênfase que ela coloca na frase seguinte: “achei um jeito”, o que novamente alerta Zoila. No questionamento da mãe, nota-se nova ocorrência do padrão ascendente do *sentence stress*, que, nesse caso, para a língua inglesa, implica que o interlocutor deverá responder de modo mais completo, como Valentina o faz, explicando a fonte do pagamento.

Na Figura 3, o conflito entre mãe e filha se instala, com ambas defendendo intensamente seus argumentos. Como se observa superficialmente, não somente o *sentence stress* ascendente é mais recorrente, como o volume das vozes e as ocorrências de *word stress*. Acreditando no bom caráter da oferta (“se oferecerem”), Valentina altera o tom de voz primeiro, gritando com a mãe “é o que sempre sonhamos

⁴ Valentina: Então, tenho boas notícias. Zoila: Bom, não guarde para si. *A ver*. Valentina: Vou começar a faculdade no outono. Zoila: Eles mudaram de ideia sobre a bolsa de estudos! Valentina: Não exatamente. Achei um jeito de pagar as mensalidades. Zoila: Como fez isso? Valentina: O Sr. Pettigrove falou comigo. Ele me falou sobre a Sra. Delatour e o pedido que ela te fez. Zoila: Ai, meu Deus. (tradução nossa).

e eu poderei fazer faculdade!”. Essa estratégia não é adotada por Zoila, quem mantém o mesmo tom de voz e faz uso de escolhas lexicais mais agressivas sobre a própria filha, usando o *word stress* para ferir a inocência da filha em “vendendo uma parte de si” e “não te criei para ser tão burra ou tão desesperada”.

Figura 3: Excerto de análise 2

00:21:26,720 - 00:21:33,150	- And <u>now</u> he's offering to pay <u>my tuition</u> if I help them. - Absolutely not.	Zoila se apoia no balcão, enquanto Valentina usa as mãos para enfatizar o que diz à mãe. Zoila levanta a mão e com um gesto circular se afasta ainda mais de Valentina, pegando outra bolsa cheia de compras no balcão.
00:21:33,780 - 00:21:38,590	Mami... this is <u>what we always dreamed of</u> . I'll <u>get to go to school!</u>	Valentina vai atrás de Zoila até o outro balcão da cozinha, com uma mão sobre o balcão realça o valor da proposta.
00:21:38,920 - 00:21:45,740	By <u>selling</u> a part of yourself? I didn't raise you to be this stupid... <u>or</u> this desperate.	Zoila usa gestos negativos de cabeça e o afastamento do tronco para expressar sua opinião.
00:21:46,810 - 00:21:48,680	I wouldn't just do this for myself. I'd be doing it for you, too.	Zoila continua retirando coisas da bolsa de compra.
00:21:51,100 - 00:21:52,500	Fof me?!	Zoila volta a encarar Valentina com a cabeça, mas não com o tronco. Mantendo a cabeça semi erguida me posição passivo agressiva.
00:21:52,750 - 00:22:01,100	I <u>see</u> the sacrifices you make, how <u>hard</u> you work on weekends and holidays, just to make a few extra bucks that you end up giving to me.	Valentina mexe a mão para enfatizar o que expressa com volume crescente de voz. Sua expressão é angustiada.
00:22:01,135 - 00:22:04,640	I don't complain about that.	Zoila dá as costas à filha e se afasta novamente de Valentina.
00:22:06,480 - 00:22:08,930	Which makes it <u>even worse!</u> <u>Now</u> ..is my chance to <u>help</u> you.	Valentina com uma mão aberta sobre o balcão realça o valor da proposta.
00:22:09,560 - 00:22:17,540	- If the tuition is paid for, you could <u>finally</u> stop and take some time off. - Do you really think that I would trade a <u>grandchild for a vacation</u> ?!	Zoila se afasta do balcão e com as duas mãos em frente ao corpo faz um gesto vertical, o qual começa com os dedos unidos embaixo e se abrem para mãos espalmadas na altura do tronco que se agitam.

Fonte: autoria própria com base em Cavalcanti (2017)⁵

No contexto das relações étnicas e econômicas da série, a fala de Zoila retoma a assimetria de poder entre os americanos ricos e brancos e os latinos pobres e assalariados, tratados como cidadãos de segunda categorias e cujos próprios corpos são usados como moeda de troca (principalmente por meio da objetificação e da sexualização exacerbada das mulheres latinas). Nesse sentido, vemos que os recursos da oralidade não funcionam somente para a contra-argumentação, mas são também estruturas ideológica do discurso⁶.

Cabe ainda comentar que a medida que a oposição se instaura entre mãe e filha, há maior deslocamento das personagens pelo cenário – fazendo com que a relação de discordância seja também representada no modo visual: Zoila evita o contato físico com a filha como forma de expressar a rejeição à sua proposta.

No final desse segmento da discussão entre Zoila e Valentina, há uma nova abordagem argumentativa de Valentina, pautada nos sacrifícios maternos, e que retomam o uso do volume crescente de voz e a ênfase em palavras que possam soar positivas para a mãe como “agora é minha chance de ajudá-la”. Todavia,

5 Valentina: E agora, se ofereceram para pagar minha mensalidade se eu os ajudar. Zoila: De modo algum. Valentina: Mami...é o que sempre sonhamos e eu poderei fazer faculdade! Zoila: Vendendo uma parte de si? Eu não te criei para ser tão burra ou desesperada. Valentina: Não faria isso só por mim. Faria isso por você, também? Zoila: por mim? Valentina: Eu vejo o sacrifício que faz, como trabalha duro nos finais de semana e nos feriados, só para fazer um extra que você acaba me dando. Zoila: Eu não reclamo disso. Valentina: O que é ainda pior! Agora, é minha chance de ajudá-la. Se a faculdade estiver paga, poderá parar e tirar uma folga. Zoila: Acha que eu trocária um neto por férias?! (tradução nossa).

6 Para uma discussão mais detalhada, ver Van Dijk, 2006.

para Zoila, a oferta da patroa não diz respeito à melhoria na sua qualidade de vida, mas uma troca entre dinheiro e uma vida. Nesse sentido, seu uso de ênfase em “neto” e “férias” e o *sentence stress* ascendente não somente passam perplexidade, mas buscam apontar as falhas de interpretação da argumentação da filha sobre o acordo proposto.

No final da cena (Figura 4), vemos que o uso de recursos da oralidade se inverte, com Valentina usando tom de voz mais estável e volume inalterado, sem ênfase em palavras específicas e Zoila mantendo a alteração do volume com *stress* em pontos específicos.

Em contrapartida, os gestos de Valentina se intensificam, enquanto os de Zoila ficam mais suaves e maternais. Com isso, as personagens mostram ao telespectador que exercem controle sobre seus corpos, evitando gesticulações e deslocamentos pelo cenário – ainda que suas vozes possam estar alteradas. De outro modo, a suavidade maternal dos gestos de Zoila associam-se a uma atitude maternalmente impositiva: ela proíbe Valentina de fazer algo.

Como se vê, foi possível nessa estrutura de transcrição e descrição, justapor diferentes recursos da oralidade que possibilitam ao leitor da análise, por meio da escrita, acessar os sentidos primeiramente expressos no material audiovisual. Com isso, é possível para a transcrição multimodal dar conta de aspectos da prosódia fundamentais para a caracterização da fala como um fenômeno não uniforme, e mostrar que os efeitos de sentido derivam tanto do que é dito quanto de como as personagens o dizem. O fato que na televisão raramente ocorre sobreposição de falas, hesitações ou reformulações também contribuiu para a transcrição.

O modo como transcrevemos os dados da oralidade faz com que os efeitos de sentido da fala não sejam vistos em isolamento, mas em conjuntura com os sentidos derivados das ações das personagens – ainda que não tenhamos recorrido a uma montagem das falas com as imagens. Além disso, através dos códigos para ênfase, hesitação e prolongamento de vogais, é possível compreender

Figura 4: Excerto de análise 3

00:22:17,610 -- 00:22:22,300	It's just an egg! And this <u>money</u> -would change-our lives.	Valentina agita energicamente o braço esticado em direção ao balcão pontuando a relevância do dinheiro oferecido, pronunciado cada expressão com dentes cerrados.
00:22:22,335 - 00:22:28,040	This <u>decision</u> would change our lives in ways you can't <u>begin</u> to understand!	Zoila se aproxima de Valentina com ambas as mãos erguidas na altura do peito e com os dedos unidos. Enquanto fala, suas mãos se abrem para abranger a filha em seu gesto.
00:22:31,110 - 00:22:35,640	I'm sorry, sweetheart. But I will not <u>allow you</u> to do this.	Zoila passa as mãos pelas têmporas e segurando os ombros de Valentina, olha-nos olhos e a proíbe de agir. Ela respira fundo, segura a cabeça da filha, toca seus ombros e se afasta.
00:22:43,430 - 00:22:48,810	- Mami, you can't stop me. - <u>What</u> did you say?	Zoila se vira do balcão para Valentina.
00:22:49,370 - 00:22:55,570	It's my body, so it's my decision. And I'm doing this for both of us, whether <u>you</u> like it or not.	Valentina dá as costas à Zoila ao final da sua fala.

Fonte: autoria própria com base em Cavalcanti (2017)⁷

⁷ Valentina: É só um óvulo! E esse dinheiro mudaria nossas vidas. Zoila: Essa decisão mudaria nossas vidas de formas que você nem sonha em entender. Desculpe, meu amor, mas não vou permitir que faça isso. Valentina: Mami, não pode me impedir. Zoila: O que você disse? Valentina: É meu corpo, portanto é minha decisão. E estou fazendo isso por nós duas, você gostando ou não.

as funções dos recursos da oralidade e suas implicações discursivas, sem desconsiderar os parâmetros contextuais como relações de poder e laços interpessoais.

Considerações Finais

A pesquisa multimodal é de natureza eclética tanto em perspectivas teóricas quanto em recursos analíticos, o que ao mesmo tempo incita produção de conhecimentos cada vez mais diversificados, interdisciplinares e aprofundados, também oferece o risco de inconsistência de dados ou, até, de confusão metodológica. Há muito a ser estabelecido tanto no que diz respeito às teorias quanto às práticas de transcrição, e análise – principalmente para *corpus* audiovisual. Sem sobrecarregar o leitor com códigos de transcrição, a proposta de análise que elaboramos com base na teoria kineicônica não oferece restrição quanto ao material audiovisual levado para análise. De modo particular, ainda, olhamos para a transcrição de dados da oralidade, que apenas simulada na série televisiva, permite dar conta de outras propostas de sentido não presentes nos estudos mais conhecidos de análise da conversação.

Nos excertos trazidos, mãe e filha que discutem a decisão de vender a si/a outro para realizar um sonho e, dando continuidade às representações midiáticas sobre a mulher latina (sensuais e capazes de usarem seu corpo para conseguirem o quiserem), para essas personagens, a série não recorre à lógica explícita da prostituição. No primeiro caso, a venda de óvulos retoma os discursos da venda do corpo e do direito sobre si, mas principalmente, do tráfico humano – acrescentado um valor criminal a suas representações.

Nosso modelo de transcrição, portanto, enfatiza que não é somente o conteúdo verbal que importa em uma análise de audiovisual, mas *como* esse conteúdo é expresso pelas personagens em conjunção com outros modos contribuintes do

formato. Isso coloca em evidência a necessidade de se levar em consideração os recursos prosódicos da oralidade, além de fenômenos da sequencialidade da interação, para compreender quais aspectos do conteúdo da fala das personagens são mais ou menos evidentes para os telespectadores e podem afetá-los de modo mais ou menos direto.

Referências Bibliográficas

BALDRY, A.P; THIBAUT, P. J. Towards Multimodal Corpora. In: Aston, G. and Burnard, L. (eds.), *Corpora in the Description and Teaching of English*. Bologna: CLUEB, 2001. pp. 87-102.

_____. *Multimodal transcription analysis: a multimedia toolkit and coursebook with associated on-line course*. Londres: Equinox Publishing, 2006.

BURN, A. *The kineikonic mode: towards a multimodal approach to moving image media*. Londres: National Centre for Research Methods Working Paper. 2013. Disponível em: <http://eprints.ncrm.ac.uk/3085/1/KINEIKONIC_MODE.pdf> acessado em: 16 ago. 2016.

BURN, Andrew. A Very Long Engagement: English and the Moving Image. In: WYSE, D., ANDREWS, R. e HOFFMAN, J. (eds) *The Handbook of English, Language and Literacy Education*. Londres: Routledge. 2010. pp 354-366. Disponível em:

<<https://aburn2012.files.wordpress.com/2014/04/burn-2009-english-and-the-moving-image.pdf>> Acesso em: 29 abril 2016

BURN, Andrew e PARKER, David. *Analyzing Media Texts*. Londres: Continuum, 2003.

JEWITT, C., BEZEMER, J. e O'HALLORAN, K. *Introducing Multimodality*. Nova York: Routledge, 2016.

CAVALCANTI, Larissa de Pinho. *Deviant maids: representações sociais sobre as mulheres latinas em uma perspectiva multimodal*. Tese de doutorado. 269 f. 2017. Recife. Universidade Federal de Pernambuco. Programa de Pós-Graduação em Linguística. Recife: UFPE, 2017.

- DIONISIO, A. P. Multimodalidade discursiva na atividade oral e escrita. In: MARCUSCHI, L. A. e DIONISIO, A. P. (org.) *Fala e escrita*. Belo Horizonte: Autêntica, 2007. pp. 177-196.
- LOPEZ, D. A Simulação da Oralidade no Telejornalismo Brasileiro. In: *Revista PJ:BR*, n.4, 2004. Disponível em: <http://www2.eca.usp.br/pjbr/arquivos/artigos4_g.htm> Acesso em 05. abril. 2018.
- MARCUSCHI, L. A. Oralidade e escrita. *Signótica*, n.9, 1997. pp. 119-145.
- MARCUSCHI, L. A. A oralidade no contexto dos usos linguísticos: caracterizando a fala. In: MARCUSCHI, L. A. e DIONISIO, A. P. (org.) *Fala e escrita*. Belo Horizonte: Autêntica, 2007. pp. 57-84.
- MARCUSCHI, L. A. e DIONISIO, A. P. Princípios gerais para o tratamento das relações entre a fala e a escrita. In: _____. (org.) *Fala e escrita*. Belo Horizonte: Autêntica, 2007.
- NORRIS, S. *Identity in Inter(action): introducing multimodal inter(action) analysis*. Berlim: De Gruyter Mouton, 2011.
- NORRIS, Sigrid. **Analyzing multimodal interaction: a methodological framework**. Londres, Routledge, 2004.
- _____. *Identity in Inter(action): introducing multimodal inter(action) analysis*. Berlim: De Gruyter Mouton, 2011.
- PEREIRA, Germana da Cruz. *As representações do gênero feminino no seriado televisivo A grande família: uma análise crítica do discurso imagético-verbal*. Universidade Federal do Ceará. Centro de Humanidades. Tese de Doutorado. 155f. Fortaleza: Ceará, 2014. Disponível em: <http://www.repositorio.ufc.br/bitstream/riufc/10509/1/2014_tese_gcpereira.pdf> Acesso em: 04 julho de 2016.
- UFRJ. *Projeto Norma Linguística Urbana Culta*. Disponível em: <<http://www.nurcrj.letas.ufrj.br/>> Acesso em: 10 ago. 2018.
- SACKS, Harvey et al. Sistemática Elementar para a organização da tomada de turnos para a conversa. In: *Veredas*, vol. 7, n. 1-2, 2003.
- VAN DIJK, Teun. Ideology and discourse analysis. In: *Journal of Political Ideologies*, vol 11(2), 2006. pp.115-140. Disponível em: <<http://www.discourses.org/OldArticles/Ideology%20and%20Discourse%20Analysis.pdf>> Acesso em: 30 julho 2015.

Submissão: 28 de agosto de 2018

Aceite: 14 de outubro de 2018.