

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS:  
ESTUDOS LINGUÍSTICOS E LITERÁRIOS

Leonice Passarella dos Reis

**WHAT DO YOU MEAN? NUCLEAR STRESS IN ENGLISH AS AN  
INTERNATIONAL LANGUAGE:  
USES AND INTERPRETATIONS**

Tese submetida ao Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários da Universidade Federal de Santa Catarina para a obtenção do Grau de Doutor em Letras.

Orientadora: Profa. Dra. Rosane Silveira

Florianópolis  
2017

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Passarella-Reis, Leonice

What do you mean? Nuclear stress in English as  
an international language : uses and  
interpretations / Leonice Passarella-Reis ;  
orientador, Rosane Silveira, 2017.

352 p.

Tese (doutorado) - Universidade Federal de Santa  
Catarina, Centro de Comunicação e Expressão,  
Programa de Pós-Graduação em Inglês: Estudos  
Linguísticos e Literários, Florianópolis, 2017.

Inclui referências.

1. Inglês: Estudos Linguísticos e Literários. 2.  
Prosody. 3. English as an international language.  
4. Nuclear stress. 5. Comunicação. I. Silveira,  
Rosane. II. Universidade Federal de Santa Catarina.  
Programa de Pós-Graduação em Inglês: Estudos  
Linguísticos e Literários. III. Título.

Leonice Passarella dos Reis

**WHAT DO YOU MEAN? NUCLEAR STRESS IN ENGLISH AS  
AN INTERNATIONAL LANGUAGE: USES AND  
INTERPRETATIONS**

Esta tese foi julgada adequada para a obtenção do Título de “Doutor em Letras” e aprovada em sua forma final pelo Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários, da Universidade Federal de Santa Catarina.

Florianópolis, 12 de junho de 2017.

---

Prof.<sup>a</sup> Anelise Reich Corseuil, Dr.<sup>a</sup>  
Coordenadora do Curso

**Banca Examinadora:**



---

Prof.<sup>a</sup> Rosane Silveira, Dr.<sup>a</sup>  
Orientadora

Universidade Federal de Santa Catarina

---

Prof.<sup>a</sup> Eva Christina Orzechowski Dias, Dr.<sup>a</sup>  
Universidad de Los Andes (Videoconferência)



---

Prof.<sup>a</sup> Hanna Kivistö-de-Souza, Dr.<sup>a</sup>  
Universidade Federal de Santa Catarina



---

Prof.<sup>a</sup> Izabel Christine Seara, Dr.<sup>a</sup>  
Universidade Federal de Santa Catarina



To those who value the benefits of  
pronunciation teaching to speakers of  
English as an international language.



## ACKNOWLEDGEMENTS

This doctoral dissertation is a miracle and one of the greatest proves of the existence of God in my life. God gave me all I needed in order to start, keep on doing it, never give up and finish it. My adviser herself is one of the angels that I have met on the way. So, I would like to express my sincere gratitude to my advisor Professor Rosane Silveira for the continuous support, caring, patience, motivation, and trust. She trusted me in moments that even I doubted myself. Professor, you are an example to be followed. You understood me right when I needed to be understood most.

I would like to thank Prof. Walcir Cardoso and Prof. Andressa Brawerman-Albini whose comments on the research project brought in valuable contributions to the design of the present study.

My thanks also go to the professors who were part of the committee: Izabel Christine Seara, Eva Christina Orzechowski Dias, and Hanna Kivistö-de-Souza. Your careful reading and suggestions have surely brought in more consistency and more value to the present research.

I would like to express my gratitude to everyone in FONAPLI for their support during data collection, especially Izabel Christine Seara, João Paulo Acosta Luz, Daise Fabiana Ribeiro Pereira Carpes, Leonan Quadros, and Karina Zendron da Cunha who even babysitted while I collected data with a mother-participant. A thank you also goes to the members of NUPFALLE Research Project for all the support and encouragement given. You all were amazing.

In addition, a thank you to Professor Juan Manoel Sosa, who introduced me to pitch analysis, taught me how to use Praat, and most importantly, to adjust the pitch settings in order to avoid the octave error and to trust my ears when *Praat* would get fuzzy.

I express my warm thanks to Alison Roberto Gonçalves for the partnership during the courses taken, his friendship and willingness to help me in whatever I needed on the course of writing this dissertation. Alison, you are such an incredible person.

This piece of research could not have been done without the willingness and dedication of the people who participated: the raters, speakers, and listeners. Every piece of data collected warmed my heart. Seeing you participating with such an open heart definitely made data collection an easy phase to go through. No words in any language will

translate how thankful I am for counting on you as participants. Thanks a zillion.

Other angels in my life were the other four members of my workplace quintet: Fernanda Moreira da Silva, Gabriela Carames Beskow, Giana Fonseca Reis, and Luci Schmoeller. Thanks for understanding my absences at the quintet's reunions. My gratitude is also extended to Captain José Luiz Bezerra Cruz, who encouraged me and did whatever was at his reach so that I could attend the course classes, lectures, and meetings with my advisor. In addition, I extend my gratitude to Chief Petty Officer Edézio who made available places where I could concentrate and work productively. Thank you.

I would like to thank Veralucia Lima dos Reis dos Santos who helped with the housework during the last three years. Lucia, a sua doação foi incrível. Você me ajudou em momentos em que você precisava de ajuda. Você esteve presente quando você não poderia estar. Esteve ao meu lado mesmo nas suas horas mais difíceis. Serei eternamente grata por isso. Muito obrigada.

I thank all my friends who participate in Grupo de Oração de Sexta for all the prayers and for listening to me every time I had something that I needed to get off my chest, for praying for me when my faith was weakened, and for being the presence of God in my life. I love you all.

My thanks also go to my husband's family: Lucival, Nazaré, Ellen, and Gerson. Thanks for all the prayers sent and for understanding my absences. Being part of this family is a blessing.

I would like to express my deepest gratitude to my parents, Izabel and Arlindo. Pai e mãe, vocês fizeram de mim uma pessoa vencedora, na simplicidade com que me educaram e com a integridade de suas ações. Se venci, foi porque tive exemplo dentro de casa. Pai, o seu positivismo e a sua perseverança para lutar e se curar me inspiraram e me tornaram mais fortes. Mãe, as suas orações incessantes certamente chegaram até mim transformadas em inspiração e força. Obrigada por me amarem. Amo vocês. A special thanks also goes to my sister Ivanice Passarella, my best friend at all times. Thanks for your home, where I wrote some pages of this dissertation. Thanks for understanding my absence in important moments of our lives. I love you. I extend my gratitude to my brother Amarildo and my sister-in-law Jussara, and my nieces Julia and also goddaughter Sara, for making all the efforts to travel and meet me in Marialva even when they knew I would not be



able to spend lots of time with them. Thanks to my cousin Marcia, my sister Rosane, my nephews Daniel and also godson Danilo, and all my other family members in Marialva who have also understood my absence at the family events along these years.

If these people mentioned above are angels in my life, one is at least a saint. That is my husband Hallthmann. Love, thanks for working hard, being strong, being considerate, being unselfish, standing by me, listening, waiting, forgiving, trying, understanding, protecting, respecting, blessing, and loving me when I would not love myself. I told you that if our marriage resisted until the end of this doctoral course, it would then last forever. Guess what? It seems it will. Your love has made me a stronger person and I will forever be grateful. I am proud to be your wife. I want and I will grow old with you. I love you.

My deepest love and gratitude go to the most beautiful people in my life: my kids Lucas and Lana. Lucas, Lana, you are so little and so huge in understanding. Every time I had to say a “no” in order to work on this dissertation instead of spending time with you broke my heart. However, every quality time we spent together gave me the strength and motivation to finish this piece of research so that we could be together more often. You have been my greatest motivation. You are everything a mom hopes for. I deeply love you.



It is a mistake to assume [that] just because a language has a conventional syntax, semantics and defined lexicon, that utterances of sentences constructed in accordance with them will be understood as intended. (Green, 1989)



## ABSTRACT

Research has revolved around the pronunciation features that are important to guarantee speech intelligibility by speakers of English as an international language (IL). A considerable body of research has addressed segments while studies investigating suprasegments are scarce. In light of the proposition that unexpected nuclear stress (NS) placement may be one of the villains for communication breakdowns among speakers in the international community (Jenkins, 2000), the present study investigated how NS placement in English by 14 Brazilian Portuguese (L1-BP) speakers of English influenced the way their intents were interpreted by 14 IL-English listeners from a variety of nationalities. Participants met in pairs (a speaker and a listener) and engaged in four pairwork tasks. Task 1 aimed at making participants more comfortable with the presence of each other. Tasks 2, 3, and 4 yielded the production of 501 audio-recorded utterances that had information being elicited, corrected, and contrasted, with NS in three different utterance positions (initial, medial, and final), and with the presence of complex words (non-challenging, short-challenging, and long-challenging), and 378 interpretations made available by checking one out of three options in the listeners' sheets. Besides the tasks, other instruments used were: (a) a questionnaire, (b) a Word Familiarity Test, (c) a Pronunciation Self-Evaluation Test, and (d) an interview with the listeners. Data analysis showed that L1-BP speakers had difficulties to convey important information by means of NS especially if the information was located in the rightmost end of the utterance, and that the perception of NS in utterance-initial position were easier to perceive. It was also found that contrasting information by means of questions was more difficult if compared to eliciting and correcting information in statements, and that NS used to signal information being corrected showed more expected perceptions and interpretations. As to the presence of complex words in the utterances, it was found that both short- and long-challenging words hindered the expected NS placement and, in terms of perception, NS in utterances with long-challenging words were misperceived more often. Finally, results showed that unexpected NS placement caused noise to communication or even led to communication breakdowns, despite the convergence strategies (Giles, H. Coupland, & J. Coupland, 1991) developed by the listeners in order to accommodate to speakers' speech. These findings support the importance of teaching BP-L1 speakers of English (a) to break speech

into meaningful units by avoiding unnecessary pauses, (b) to use NS in order to signal important information (narrow-focus contexts), (c) to destress unimportant information, and (d) to place word stress in English long words.

**Keywords:** Prosody. English as an International Language. Nuclear Stress. Production. Perception. Communication.

Number of pages: 264 (excluding appendices) and 352 (including appendices)

Number of words: 67,984 (excluding appendices)

## RESUMO

Pesquisadores têm investigado quais aspectos da pronúncia são importantes para garantir um discurso inteligível por parte de falantes de inglês como língua internacional. Um corpo considerável de estudos tem investigado os segmentos enquanto estudos investigando os suprasegmentos são escassos. Tendo em vista a assertiva de que o uso inesperado do acento nuclear é um dos grandes vilões para falhas em comunicação entre falantes da comunidade internacional (Jenkins, 2000), este estudo investigou como o uso do acento nuclear em inglês de 14 falantes brasileiros de inglês influenciou o modo como suas intenções foram interpretadas por 14 ouvintes de inglês como língua internacional de diferentes nacionalidades. Os participantes se encontraram em duplas (um falante e um ouvinte) e realizaram quatro tarefas. A Tarefa 1 tinha como objetivo fazer com que os participantes se sentissem mais confortáveis na presença um do outro. As Tarefas 2, 3 e 4 deram origem à produção de 501 enunciados, gravados em áudio, com informações sendo buscadas, corrigidas e contrastadas, com o acento nuclear em três locais diferentes no enunciado (início, meio, final) e com a presença de palavras complexas (simples, complexas curtas e complexas longas). Essas tarefas também originaram 378 interpretações feitas através da escolha de uma das três opções para cada enunciado apresentadas nas folhas dos ouvintes. Além das tarefas, outros instrumentos foram utilizados: (a) um questionário, (b) um Teste de Familiaridade com o Vocabulário, (c) um Teste de Auto-Avaliação da Pronúncia e (d) uma entrevista com os ouvintes. A análise dos dados demonstrou que os falantes brasileiros tiveram dificuldades para transmitir informações importantes por meio do acento nuclear, principalmente se a informação estivesse localizada ao final do enunciado, e que o acento nuclear em posição inicial foi percebido com mais facilidade do que quando em outras posições. Também foi evidenciado que contrastar informações em enunciados interrogativos foi mais difícil do que fornecer e corrigir informações em declarativas. Com relação à presença de palavras complexas, constatou-se que tanto palavras complexas curtas quanto palavras complexas longas prejudicaram o uso esperado do acento nuclear e que, em termos de percepção, o acento nuclear em enunciados com palavras complexas longas foram mais frequentemente não percebidos. Finalmente, os resultados indicaram que o uso inesperado do acento nuclear causou ruídos à comunicação ou, até mesmo, levou a falhas na comunicação,

apesar de os ouvintes terem desenvolvido estratégias de convergência (Giles, H. Coupland & J. Coupland, 1991) para se acomodarem ao discurso dos falantes. Esses resultados destacam a importância de se ensinar, aos falantes brasileiros de inglês, como (a) quebrar o discurso em unidades significativas, evitando-se pausas desnecessárias, (b) usar o acento nuclear para sinalizar a informação importante no enunciado (em contextos de foco estreito), (c) retirar a proeminência de informações que não sejam importantes e como (d) atribuir o acento lexical na sílaba adequada em palavras longas em inglês.

**Palavras-chave:** Prosódia. Inglês como Língua Internacional. Acento Nuclear. Produção. Percepção. Comunicação.

Número de páginas: 264 (excluindo apêndices) e 352 (incluindo apêndices)

Número de palavras: 67.984 (excluindo apêndices)



## LIST OF FIGURES

Figure 1:	<i>Legumes are a good source of VITAMINS.</i> Production with the NS on vitamins, with a falling-rising intonational pattern .....	59
Figure 2:	<i>Are legumes a good source of VITAMINS?</i> Production with the NS on vitamins, with a rising intonational pattern.....	60
Figure 3:	<i>Legumes are a GOOD source of vitamins.</i> Production with the nuclear stress on GOOD with a falling-rising pattern .....	61
Figure 4:	<i>Are legumes a GOOD source of vitamins?</i> Production with the NS on GOOD with a rising pattern .....	61
Figure 5:	<i>Your FATHER won't be able to HELP me.</i> .....	73
Figure 6:	<i>CLARO que é do Guilherme "OF COURSE it is Guilherme's</i> produced by a speaker from Minas Gerais (Brazil). .....	74
Figure 7:	<i>Is ZENNY afraid of heights or is it ANA?</i> – An example of a wide-scope disjunctive question with closed pronunciation.....	76
Figure 8:	<i>O professor de MATEMÁTICA foi visitar a sobrinha?</i> (Has the mathematics teacher visited his niece?) .....	78
Figure 9:	Random pitch settings. ....	133
Figure 10:	Pitch values set near the highest and the lowest limits of the curves. ....	133
Figure 11:	Praat window displaying vertical scales transformed from Hertz to semitones. ....	135
Figure 12a:	Production of <i>The T-shirt shrank A LOT</i> by speaker B-01, an illustration of an utterance with more than one thought group caused by pauses.....	137
Figure 12b:	Production of <i>The T-shirt shrank A LOT</i> by speaker B-01, an illustration of wave portions that were not pitch tracked. ....	138

Figure 13: <i>ZENNY loves you</i> (Task 2, Item 04) by B13 – An example of an expected NS placement.....	154
Figure 14: <i>Does Zenny LIKE shrimp?</i> (Task 4, Item 09) by B09 – An example of a NS placed on a position other than the expected. ....	155
Figure 15: <i>No, the rhythm suits ZENNY</i> (Task 3, Item 03) by B03 – An example of NS placed in more than one location with a greater pitch change range on the expected portion of the utterance. ....	156
Figure 16: <i>No, MARIA had it curled</i> (Task 3, Item 09) by B09 – An example of NS placed in more than one location with no distinction in pitch change ranges. ....	157
Figure 17: <i>No, the rhythm suits ZENNY</i> (Task 3, Item 03) by B02 – An example of NS placed in more than one location with a greater pitch change at an unexpected portion of the utterance. ....	159
Figure 18: <i>No, Ana LOVES dogs</i> (Task 3, Item 07) by B09 – An example of utterances in which pitch changes did not reach the three-semitone-pitch change. ....	160
Figure 13: <i>ZENNY loves you</i> (Task 2, Item 04) by B13 – An example of an expected NS placement.....	170
Figure 19: <i>The boss LOVES your world</i> (Task 2, Item 08) by B07 – An example of an expected production of utterances with short-challenging words. ....	171
Figure 20: <i>Zenny DISLIKES vinegar</i> (Task 2, Item 03) by B14 – An example of an expected production of utterances with long-challenging words. ....	172
Figure 21: <i>Zenny LOVES you</i> (Task 2, Item 04) by B01 – An example of an unexpected production of utterances with non-challenging words. ....	173
Figure 22: <i>The boss LOVES your world</i> (Task 2, Item 08) by B06 – An example of an unexpected production of utterances with short-challenging words. ....	174

Figure 23: <i>The government talked to ZENNY</i> (Task 2, Item 06) by B14 – An example of an unexpected production of utterances with long-challenging words. ....	175
Figure 24: <i>Did Ana twirl her HAIR?</i> (Task 4, Item 07) repeated by B12 – An example of an unexpected perception caused by unexpected pronunciation as reported by the listener. ....	188
Figure 25: <i>Did Ana twirl her HAIR?</i> (Task 4, Item 07) by B11 – An example of listeners’ guiding their decisions on an aspect other than NS placement. ....	189
Figure 26: <i>The government talked to ZENNY</i> (Task 2, Item 06) by B03 – An example of perceived pauses hindering the expected interpretation of speakers’ intent. ....	191
Figure 27: <i>Is the atmosphere around Zenny GOOD?</i> (Task 4, Item 05) by B10 – An example of listeners’ uncertainty of final pitch change signaling important information or the pattern of a question. ....	192
Figure 28: <i>ZENNY traveled to Pennsylvania</i> (Task 2, Item 03) by B12 – An example of an expected NS placement that had sites of reduced interpretability, but was interpreted as expected (first-time production).....	200
Figure 29: <i>ZENNY traveled to Pennsylvania</i> (Task 2, Item 03) by B12 – An example of an expected NS placement that had sites of reduced interpretability, but was interpreted as expected (repetition). ....	200
Figure 30: <i>The rhythm suits ZENNY</i> (Task 3, Item 03) by B09 – An example of a NS assigned to an unexpected position perceived at the expected position. ....	203
Figure 23: <i>The government talked to ZENNY</i> (Task 2, Item 06) by B14 – An example of an unexpected production of utterances with long-challenging words. ....	206
Figure 31: <i>Did Ana twirl her HAIR?</i> (Task 4, Item 07) by B03 – An example of a communication breakdown resulted from an utterance said with two or more NSs, with a greater pitch change range on the expected portion. ....	207

Figure 32: <i>Unfortunately Zenny AGREED</i> (Task 3, Item 05) by B06 – An example of an expected perception of NS in an utterance said with two or more NSs, with no distinct pitch change ranges. ....	208
Figure 33: <i>MARIA had it curled</i> (Task 3, Item 09) by B13 – An example of a communication breakdown caused by the convergence strategy that pitch change in initial position was a trait of the speakers’ speech rather than the portion with important information. ....	210
Figure 34: <i>Is the atmosphere around Zenny GOOD?</i> (Task 4, Item 05) by B03 – An example of a communication breakdown caused by the lack of convergence to the fact that pitch change in initial position was a trait of the speakers’ speech rather than the portion with important information. ....	211
Figure T20: <i>Did Ana twirl her HAIR?</i> (Task 4, Item 26) by B14 – An example of an unexpected NS production combined with unexpected syllable lengthening/rising intonation, unexpected pauses, and unexpected pronunciation. ....	217
Figure T5: <i>No, MARIA had it curled</i> (Task 3, Item 09) by B12 – An example of an expected NS production with pronunciation issues in utterances with short-challenging words. ....	223
Figure 35: <i>MARIA had it curled</i> (Task 3, Item 09) repeated by B12. ....	224
Figure 36: <i>Ana loves ZENNY</i> (Task 2, Item 05) by B01 – An example of unexpected pronunciations of non-challenging words affecting listeners’ perception. ....	226
Figure 37: <i>Ana loves ZENNY</i> (Task 2, Item 05) repeated by B01. ....	227
Figure 38: <i>Does Zenny LIKE shrimp?</i> (Task 4, Item 09) by B01 – An example of unexpected pronunciations of non-challenging words contributing to listeners’ perception of NS. ....	228
Figure R1: Production of <i>Ana loves ZENNY</i> (Task 2) by speaker B07 – An example of different pitch-track tracing by Praat. ....	316

Figure R2: Production of <i>Zenny DISLIKES vinegar</i> (Task 2, Item 3), produced by speaker B09 – An example of different pitch-track tracing by Praat. ....	317
Figure R3: Production of <i>The T-shirt shrank A LOT</i> (Task 2, Item 10) by B03 – An example of confusing final pitch change.....	318
Figure S1: <i>Zenny loves DOGS</i> (Task 3, Item 04) by B05 – An example of a dismissed case (1). ....	319
Figure S2: <i>Do MEN lie a lot?</i> (Task 4, Item 04) by B1 – An example of a dismissed case (2). ....	320
Figure T1: <i>ZENNY loves you</i> (Task 2, Item 04) by B13 – An example of an expected NS placement in utterances with non-challenging words. ....	324
Figure T2: <i>The boss LOVES your world</i> (Task 2, Item 08) by B07 – An example of an expected NS production with no pronunciation issues in utterances with short-challenging words.....	325
Figure T3: <i>Zenny DISLIKES vinegar</i> (Task 2, Item 03) by B14 – An example of an expected NS production with no pronunciation issues in utterances with short-challenging words.....	326
Figure T4: <i>No, Ana LOVES dogs</i> (Task 2, Item 07) by B03 – An example of an expected NS production with pronunciation issues in utterances with non-challenging words.....	327
Figure T5: <i>No, MARIA had it curled</i> (Task 3, Item 09) by B12 – An example of an expected NS production with pronunciation issues in utterances with short-challenging words.....	328
Figure T6: <i>No, RED is an appetite stimulator</i> (Task 3, Item 02) by B14 – An example of an expected NS production with pronunciation issues in utterances with long-challenging words.....	329
Figure T7: <i>ANA sent the e-mail</i> (Task 2, Item 06) by B13 – An example of an unexpected NS production with no	

pronunciation issues in utterances with non-challenging words. ....	331
Figure T8: <i>The boss LOVES your world</i> (Task 2, Item 08) by B15 – An example of an unexpected NS production with no pronunciation issues in utterances with short-challenging words. ....	332
Figure T9: <i>Is ZENNY leaving immediately?</i> (Task 4, Item 08) by B15 – An example of an unexpected NS production with no pronunciation issues in utterances with long-challenging words. ....	333
Figure T10: <i>Zenny LOVES you</i> (Task 2, Item 04) by B01 – An example of a production with both unexpected NS placement and unexpected pronunciation in utterances with non-challenging words. ....	334
Figure T11: <i>The boss LOVES your world</i> (Task 2, Item 08) by B06 – An example of a production with both unexpected NS placement and unexpected pronunciation in utterances with short-challenging words. ....	335
Figure T12: <i>The government talked to ZENNY</i> (Task 2, Item 06) by B14 – An example of a production with both unexpected NS placement and unexpected pronunciation in utterances with long-challenging words. ....	336
Figure T13: <i>The boss LOVES your world</i> (Task 2, Item 08) by B14 – An example of an unexpected production with unexpected pauses. ....	337
Figure T14: <i>Yellow STIMULATES optimism</i> (Task 2, Item 10) by B03 – An example of an unexpected NS production with both unexpected pronunciation and unexpected pauses. ....	338
Figure T15: <i>Zenny sent you the pearls</i> (Task 2, Item 02) by B01 – An example of an unexpected NS production combined with unexpected repetition and unexpected pronunciation. ....	340

Figure T16: <i>Does the meeting START at ten?</i> (Task 4, Item 06) by B05 – An example of an unexpected NS production combined with unexpected repetition in an attempt to correct NS placement. ....	341
Figure T17: <i>Is ZENNY leaving immediately?</i> (Task 4, Item 21) by B7 – An example of an unexpected NS production combined with unexpected syllable lengthening and rising intonation.....	342
Figure T18: <i>Did ANA twirl her HAIR?</i> (Task 4, Item 07) by B06 – An example of an unexpected NS production combined with unexpected syllable lengthening and pauses (with expected pronunciation). ....	343
Figure T19: <i>Unfortunately Zenny AGREED</i> (Task 2, Item 05) by B14 – An example of an unexpected NS production combined with unexpected syllable lengthening/rising intonation and unexpected pronunciation.....	345
Figure T20: <i>Did Ana twirl her HAIR?</i> (Task 4, Item 26) by B14 – An example of an unexpected NS production combined with unexpected syllable lengthening/rising intonation, unexpected pauses, and unexpected pronunciation. ....	346
Figure T21: <i>RED is an appetite stimulator</i> (Task 3, Item 2) by B13 – An example of an unexpected NS production combined with unexpected rising intonation, unexpected pauses, unexpected pronunciation, and repetition.....	347





## LIST OF TABLES

Table 1:	Combinations for speaker’s utterance intelligibility and effectiveness .....	52
Table 2:	NS in English and BP – a comparison.....	79
Table 3:	Cronbach’s Alpha Output from the Reliability Analysis..	105
Table 4:	Inter-rater correlation matrix – Reliability Analysis .....	105
Table 5:	Rater-total statistics – Reliability Analysis.....	106
Table 6:	Oral Proficiency level of participants – speakers and listeners.....	107
Table 7:	Speakers’ background data.....	110
Table 8:	Listeners’ background data.....	112
Table 9:	Pairs for the interactions for data collection.....	116
Table 10a:	Tasks, roles of speakers and listeners, and respective Appendices for the tasks.....	118
Table 10b:	Equipment used during data collection .....	119
Table 11:	Sequences of Tasks 2, 3, and 4.....	130
Table 12a:	Dismissed cases.....	140
Table 12b:	Summary of the variables and their levels according to each research question.....	141
Table 13:	NS placement by Brazilians in English – general results .	147
Table 14:	Production results according to NS location – first-time production.....	148
Table 15:	Production results according to NS location - repetition..	148
Table 16:	Production results rearranged for statistical testing (2 x 3 Contingency Table) – IV Position .....	149
Table 17:	Production results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Position .....	149
Table 18:	Production details according to NS location – first-time production.....	152

Table 19:	Production details according to NS location – repetition.	153
Table 20:	Production results according to type of information – first-time production.....	162
Table 21:	Production results according to type of information – repetition .....	162
Table 22:	Production results rearranged for statistical testing (2 x 3 Contingency Table) – IV Type of information.....	162
Table 23:	Production results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Type of Information .....	163
Table 24:	Utterances produced in Tasks 2, 3, and 4.....	166
Table 25:	Pronunciation Self-Evaluation Test - Speakers.....	167
Table 26:	Production results according to complexity of words – first-time production.....	169
Table 27:	Production results according to complexity of words – repetition .....	169
Table 28:	Production results rearranged for statistical testing (2 x 3 Contingency Table) – IV Complexity of words.....	176
Table 29:	Production results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Complexity of words.....	176
Table 30:	Number of aspects in unexpected productions: a comparison.....	178
Table 31:	General perception results.....	181
Table 32:	Unexpected perceptions: Repetition meaning reduced interpretability only vs repetition leading to real communication breakdowns.....	182
Table 33:	Unexpected perceptions: real communication breakdowns – based on repetition Vs based on first production .....	182
Table 34:	Production and perception: a comparison of frequencies .....	183
Table 35:	Listeners’ opinions on why they interpreted the speakers’ intent as expected.....	185

Table 36:	Listeners' opinions on why they did not interpret the speakers' intent as expected. ....	186
Table 37:	Perception results according to expected NS position.....	195
Table 38:	Perception results rearranged for statistical testing (2 x 3 Contingency Table) – IV NS Location.....	196
Table 39:	Perception results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV NS Location.....	196
Table 40:	Perception of expected and unexpected productions - Location.....	197
Table 41:	Perception and production details according to NS position in the utterances – expected allocations.....	199
Table 42:	Perception and production details according to NS position in the utterances – unexpected allocations (Case 01).....	202
Table 43:	Perception and production details according to NS position in the utterances – placement of NS in two or more locations or nowhere – Cases 02, 03, 04, and 05 ....	204
Table 44:	Perception according to the unexpected productions of NS detailed according to position – general.....	212
Table 45:	Perception results according to type of information.....	214
Table 46:	Perception results rearranged for statistical testing (2 x 3 Contingency Table) – IV Type of information.....	215
Table 47:	Perception results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Type of Information.....	215
Table 48:	Unexpected perception results according to type of information – Repetition meaning reduced interpretability only vs repetition leading to real communication breakdowns.....	217
Table 49:	Unexpected perception results according to type of information – real communication breakdowns based on repetition Vs based on first production.....	218
Table 50:	Perception results according to type of information and production status.....	219

Table 51: Perception results according to the complexity level of words in the utterances.....	221
Table 52: Perception results rearranged for statistical testing (2 x 3 Contingency Table) – IV Complexity of words.....	222
Table 53: Perception results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Complexity of words.....	222
Table 54: Unexpected perception results according to complexity of words – Repetition meaning reduced interpretability only vs repetition leading to real communication breakdowns .....	225
Table 55: Unexpected perception results according to complexity of words – real communication breakdowns based on repetition Vs based on first production .....	225
Table 56: A comparison between production and perception – complexity of words.....	230
Table 57: Total counts of expected productions and expected perceptions for each utterance.....	231
Table 58: Number of aspects in unexpected productions and perception: a comparison .....	233
Table T1: Pronunciation details – first-time production.....	323
Table U1: Pronunciation details and perception .....	350

## LIST OF ABBREVIATIONS AND SYMBOLS

AL:	Additional Language
BP:	Brazilian Portuguese
CAT:	Communication Accommodation Theory
FL:	Foreign Language
IL:	International Language
IL-English speakers:	speakers of English as an international language
L1:	First Language
L1-BP speakers:	speakers of Brazilian-Portuguese as a first language
L2:	Second Language
L1-English speakers:	speakers of English as a first language
ms:	millisecond(s)
NS:	nuclear stress
s:	second(s)
st:	semitone(s)
%:	thought group boundary
L:	low
H:	high



## TABLE OF CONTENTS

<b>1</b>	<b>INTRODUCTION.....</b>	<b>35</b>
1.1	Context of Investigation and Statement of the Problem.....	35
1.2	Statement of purpose.....	37
1.3	Relevance of the study .....	37
1.4	Research Questions and Hypotheses.....	38
1.5	Organization of the dissertation .....	40
<b>2</b>	<b>REVIEW OF LITERATURE .....</b>	<b>43</b>
2.1	Language, communication, and nuclear stress.....	43
2.1.1	Language.....	43
2.1.2	English language status and speakers of English .....	45
2.1.3	Communication, Interaction, and Communication Accommodation Theory.....	48
2.1.4	Speech Communication, Perception, and Intelligibility.....	51
2.1.5	Nuclear Stress .....	55
2.1.5.1	The function of nuclear stress .....	62
2.1.5.2	Nuclear Stress in English and in Brazilian Portuguese .....	68
2.1.5.2.1	NS in statements.....	71
2.1.5.2.2	NS in disjunctive questions .....	75
2.1.5.3	The unexpected allocation of nuclear stress .....	80
2.1.6	Summary .....	82
2.2	The perception of NS and the interpretation of speakers' intent .....	83
2.2.1	Cues for NS perception .....	83
2.2.2	Production and perception of NS and speakers' intent interpretation worldwide .....	86
2.2.3	Production and perception of NS and speakers' intent interpretation in Brazil .....	95

2.2.4	Summary.....	97
2.3	Conclusion of the chapter .....	98
<b>3</b>	<b>METHOD .....</b>	<b>101</b>
3.1	Recruiting and selecting participants .....	101
3.1.1	Data collection for level of proficiency .....	102
3.1.2	Data analysis for level of proficiency .....	103
3.1.3	Results for level of proficiency.....	105
3.1.4	Summary.....	108
3.2	Participants .....	108
3.2.1	The Speakers.....	108
3.2.2	The Listeners .....	111
3.2.3	Summary.....	114
3.3	Data Collection .....	115
3.3.1	Instruments and procedures .....	115
3.3.2	Summary.....	131
3.4	Data Analysis.....	132
3.5.1	Statistics.....	141
3.5.2	Summary.....	143
3.6	Conclusion of the chapter .....	143
<b>4</b>	<b>RESULTS AND DISCUSSION .....</b>	<b>145</b>
4.1	Brazilians' placement of NS in English.....	146
4.1.1	RQ1 – Production results for location of NS .....	148
4.1.2	RQ2 – Production results according to type of information .....	161
4.1.3	RQ3 – Production results according to complexity of words.....	165
4.1.4	Summary of production results.....	179



4.2	Perception of NS and interpretation of Brazilians' intent ...	180
4.2.1	RQ4 – The interpretation of Speakers' intent – Expected vs unexpected.....	183
4.2.2	RQ5 – Perception results for the interpretation of speakers' intent according to NS position.....	194
4.2.2.1	Answering RQ5.....	195
4.2.2.2	Perception results according to actual NS locations as placed by L1-BP speakers.....	197
4.2.3	RQ6 – NS Perception and speakers' intent interpretation according to type of information.....	213
4.2.4	RQ7 – NS Perception and speakers' intent interpretation according to level of complexity of words.....	220
4.2.4.1	Answering RQ7.....	221
4.2.4.2	Perception results according to actual NS productions and complexity of words.....	229
4.2.5	Summary of perception results.....	235
4.3	Conclusion of the chapter.....	237
<b>5</b>	<b>FINAL REMARKS.....</b>	<b>239</b>
5.1	Pedagogical implications .....	241
5.2	Limitations and suggestions for future research.....	243
5.3	Concluding words .....	245
	<b>REFERENCES.....</b>	<b>247</b>
	<b>APPENDICES .....</b>	<b>265</b>
	Appendix A: Invitation do participate addressed to Brazilian Speakers of English.....	267
	Appendix B: Invitation do participate addressed to Speakers of English from a nationality other than Brazilian .....	268
	Appendix C: Consent Term - Speakers.....	269
	Appendix D: Consent Term - Listeners .....	272

Appendix E: Questionnaire.....	275
Appendix F: Oral Production Task.....	276
Appendix G: Instructions for the raters.....	285
Appendix H: Forms for the raters.....	290
Appendix I: Task 1 – Breaking the ice.....	297
Appendix J: Task 2 – I know the answer! (Speaker’s version).....	298
Appendix K: Task 2 – I don’t know the answer, do you! (Listener’s version).....	300
Appendix L: Task 3 – You’re wrong... (Speaker’s version).....	302
Appendix M: Task 3 – You’re wrong... (Listener’s version).....	304
Appendix N: Task 4 – I know what you’re saying next! (Speaker’s version).....	306
Appendix O: Task 4 – I know what you’re saying next! (Listener’s version).....	308
Appendix P: Word Familiarity Test.....	310
Appendix Q: Pronunciation Self-Evaluation Test.....	313
Appendix R: Praat pitch tracking issues.....	316
Appendix S: Dismissed Cases.....	319
Appendix T: Complexity of words and NS placement – more illustrations and some qualitative analysis.....	322
Appendix U: Complexity of words and NS perception – perception according to the production aspects.....	350

# 1 INTRODUCTION

## 1.1 CONTEXT OF INVESTIGATION AND STATEMENT OF THE PROBLEM

I know that you believe that you understood what you think I said, but I am not sure you realize that what you heard is not what I meant. (Robert McCloskey)

The quote above depicts a common phenomenon in daily life: that of misunderstandings. Daily, in different contexts, be it in family or social contexts such as school and work, communication acts continuously take place. They are essential for all areas in all parts of life. Being communication so important for human beings, human beings themselves are one of the most important agents carrying manifold factors that have an impact in communication. Among these factors is accent.

Accent<sup>1</sup> is ubiquitous. Every language, irrespective of its status being first (L1), second (L2), or foreign (FL), has variation of the pronunciation of speech sounds and prosodic features (e.g., intonation, rhythm, and stress) that may interfere with the way people produce and process information. This variation contributes to rendering the communication process more fragile.

People usually go through communication misunderstandings even in their L1, given that interlocutors involved in communication have their own expectations through which speech may be filtered. On the one hand, listeners may apply to speech an unintended meaning. On the other, depending on the choices of the speakers, they may also apply to speech unexpected meaning. In other words, “meaning lies in the ear of the beholder as much as in the mouth of the speaker” (Jenkins, 2000, p. 177). One may choose to communicate an idea by using specific linguistic structures based on past experiences, but the hearer listens to what someone says using the hearer’s own past experiences, both hearer and speaker currently undergoing a unique situation (that might resemble others, but is unique) (Beckner et al., 2009).

---

<sup>1</sup> Accent refers to the pronunciation of speech sounds (segments) and suprasegmental features (intonation, rhythm, stress, segmental length, tempo, loudness) of a given language variety (Moyer, 2013).

In international communication contexts, there is greater frangibility in communication, given that speakers and listeners from different L1 backgrounds have to deal with foreign accents, which vary according to manifold individual and sociocultural factors and the speakers' L1 itself. Foreign accents may have a great deal of undesirable consequences for international-language<sup>2</sup>-English (IL-English) speakers given that they may (1) make IL speakers' speech difficult to understand, (2) cause listeners to misjudge an IL speaker's affective state, and (3) cause negative personal evaluations (Derwing & Munro, 2005; Flege, 1995; Hahn, 2004; Lindemann, 2010; Moyer, 2013; Munro & Bohn, 2007; Zoghbor, 2010).

IL-English research has been carried out to understand the impact of IL-English pronunciation into communication mainly between members of the international community (including or not those who are L1-English speakers). One famous study in the area is that conducted by Jennifer Jenkins, who set up the *Lingua Franca Core* (LFC). According to the LFC (Jenkins, 2000), the unexpected production of specific segments and suprasegments may hinder communication to a great extent. Among the suprasegments, the scholar highlights the role played by nuclear stress (NS) placement. In short, NS signals, by means of pitch change, duration and loudness, the most important information in an utterance. The idea behind the importance of NS is that by having speech uttered as expected in segmental terms does not guarantee successful communication if not combined with the expected<sup>3</sup> placement of NS. Drawing on the idea that the relevant features for successful communication are L1 dependent and thus should be considered in a specific community of speakers (Berns, 2008), many follow-up studies have been being conducted since then in order to gain a deeper comprehension of whether or not the LFC applies to different language communities. In Brazil, these studies do not abound. Most of the ones available investigate the segmental level (e.g., Cruz, 2004)

---

<sup>2</sup> English as an International Language in the present study is an umbrella term, partly inspired in McKay (2010), that permeates three notions: (1) that of number (additional, second, third, and so forth in opposition to first language); (2) that of use (international – not favoring a specific English as an L1 variety); and (3) that of acquisition (interlanguage: it is shaped according to the L1 attractors and language use).

<sup>3</sup> *Expected* here is a term used to refer to a correspondence with a given discursive context. If NS is placed according to the discursive context, then it has an expected placement.

while studies investigating the prosodic level in terms of intelligibility are rare.

## 1.2 STATEMENT OF PURPOSE

The present study aims at contributing to the body of research that is dedicated to investigating the effect of unexpected prosody in international communication, more specifically, when used by L1-Brazilian-Portuguese (L1-BP) speakers of English. It investigates the use of NS by L1-BP speakers of English as the only cue to signal the most important information in an utterance. It also investigates the perception of NS placed by L1-BP speakers of English by members of the international community and the resulting interpretation of these speakers' intent.

A decision to include three independent variables was made. First, three types of information were chosen: being elicited, being corrected, and being contrasted. Second, the present study also addresses the effects of NS position (initial, medial, final). Finally, the role played by level of complexity of words, in terms of pronunciation, in the utterance was also investigated.

## 1.3 RELEVANCE OF THE STUDY

Unexpected NS placement may cause noise to communication or even lead to communication breakdowns. Reaching a comprehension of how NS is placed by L1-BP speakers of English, how this placement is perceived and how the consequent speakers' intent is interpreted by the members of the international community is considerably important.

The present study uses a method that triggers the full engagement of both speakers and listeners into the tasks designed for data collection. This engagement resembles that of people involved in real communication, where communicating information is the ultimate goal. As a result, although data collection involves reading, the speaker is more concerned with the message than with the decoding of the utterances. This is extremely important to prosody research and the present study contributes with the area of L2 prosody research by proposing a set of tasks to elicit, correct, and contrast information in face-to-face interactions.

The three types of information investigated pertain many communication acts in daily life, as people are often providing information, making comments for correcting mistaken information or drawing contrasts and comparisons. Reaching some understanding of how Brazilians convey these types of information and how they are interpreted while doing so is a must.

The canonical position for NS is the rightmost end of the utterance (Chomsky & Halle, 1968). Investigating how speakers and listeners deal with important information when it is not at the rightmost position is of great significance for international communication and to inform pedagogy.

Finally, IL-English speakers face challenges in terms of pronunciation both at segmental and prosodic levels. Utterances were embedded with non-challenging and challenging words in terms of pronunciation in order to understand the extent to which the presence of words difficult to pronounce interferes with the expected assignment of NS. This may promote a discussion on the importance of dealing with pronunciation issues more closely in English classes.

#### 1.4 RESEARCH QUESTIONS AND HYPOTHESES

This study aims at investigating (1) the NS assignment by L1-BP speakers of English when eliciting, contrasting, and correcting information during interactions with listeners from nationalities other than Brazilian and (2) the way these Brazilians' intent is interpreted by these listeners. It is not the aim to discuss the differences in prominence and describe the intonational patterns (the tunes – High or Low), but rather understand if the locations that are chosen by the speakers to place NS in order to signal information being elicited, corrected, and contrasted lead listeners to expected interpretations, and thus yield successful communication. The research questions and hypotheses guiding this study are the following:

RQ1: How is NS placed by L1-BP speakers of English, according to the expected position of NS (initial, medial, final) in utterances?

- H1: Based on the findings in the pilot study (Passarella-Reis & Silveira, 2016)<sup>4</sup>, Hypothesis 1 predicts that the position of NS in the utterance will not interfere with either the expected or unexpected production of NS.
- RQ2: How is NS placed by L1-BP speakers of English, according to the type of information (being elicited, being corrected, being contrasted) in utterances?
- H2: Signaling eliciting information will pose more difficulties than contrasting and correcting information. These predictions are based on the assumption that contrasting and corrective information naturally require more prominence than providing eliciting information does (Frota & Moraes, 2016; Klok et al, 2011).
- RQ3: How is NS placed by Brazilian speakers of English, according to the complexity of words (non-challenging, short challenging, long challenging) in the utterances?
- H3: Based on the findings in the pilot study, Hypothesis 3 predicts that NS will be assigned as expected less frequently when in utterances with challenging words (either short or long) if compared to utterances with non-challenging words.
- RQ4: How is the speakers' intent interpreted by IL-English listeners according to NS placement?
- H4: Unexpected NS placement will hinder the interpretation of speakers' intent while expected NS placement will not (Atechi, 1994; Jenkins, 1997; Lanham, 1984; Tiffen, 1974).
- RQ5: How is the speakers' intent interpretation affected by the position of NS (initial, medial, final) in utterances?
- H5: Utterances with NS in final position will yield more unexpected perceptions than will those with NS in initial and medial position, because the final position is the

---

<sup>4</sup> In the pilot study, participants had difficulties to place the NS irrespective of utterance position. However, words that showed to be difficult to pronounce showed to interfere.

canonical place for NS and narrow-focus NS has to increase greatly in order to be perceived in final position if compared to the other utterance positions (Calhoun, 2007).

- RQ6: How is the speakers' intent interpretation affected by the type of information (being elicited, being corrected, being contrasted) in utterances?
- H6: Based on the findings in the pilot study<sup>5</sup> and on the assumption stated in Hypothesis 2, Hypothesis 6 predicts that NS for signaling information being corrected and contrasted will yield more expected interpretations than NS for signaling information being elicited will.
- RQ7: How is the speakers' intent interpretation affected by the complexity of words (non-challenging, short challenging, long challenging) in the utterances?
- H7: Based on the findings in the pilot study, complexity level of words in utterances (non-challenging, short challenging, and long challenging) will play a role in the way listeners interpret speakers' intent, as a result of the effect it will have in speakers' production, as predicted in Hypothesis 3. Utterances that contain challenging words, either short or long, will be more frequently misinterpreted than those that contain non-challenging words.

## 1.5 ORGANIZATION OF THE DISSERTATION

Chapter 2 presents important concepts for the present study. It discusses the way language is seen in the study (Section 2.1.1), the status of English and speakers of English (Section 2.1.2), it approaches the Communication Accommodation Theory (Section 2.1.3), the concepts of intelligibility, comprehensibility and interpretability (Section 2.1.4), and it defines nuclear stress (Section 2.1.5). Finally, it reviews some studies on interpretability worldwide (Section 2.2.1) and in Brazil (Section 2.2.2).

---

<sup>5</sup> In the pilot study, it was found that misinterpretations were more frequent for unexpected productions of NS to provide information being elicited.



Chapter 3 is dedicated to the method. It introduces the procedures for selecting the participants (Sections 3.1.1 and 3.1.2), it describes the participants (Section 3.2), the instruments and procedures for data collection (Section 3.3), and the steps for data analysis (Section 3.4).

Chapter 4 reports and discusses the results obtained for the production of NS according to NS position (Section 4.1.1), types of information (Section 4.1.2), and level of difficulty of words (Section 4.1.3), and for the perception of NS and the interpretation of speakers' intent according to the speakers' productions (Section 4.2.1) and to the same three independent variables (Sections 4.2.2, 4.2.3, and 4.2.4).

Chapter 5 closes the study. It reviews the conclusions that emerged from the results discussed in Chapter 4, addresses the pedagogical implications of the findings (Section 5.1), lists the limitations of the present study, gives suggestions for future research (Section 5.2) and adds some concluding words (Section 5.3). Chapters 2, 3, and 4 have a closing section to conclude the chapters (Sections 2.3, 3.6, and 4.3).



## 2 REVIEW OF LITERATURE

This chapter is divided into three sections. Section 1 addresses the most important concepts and tackles important issues for the present study, namely, (1) language, (2) the status and the speakers of English; (3) communication, (4) interaction, (5) Communication Accommodation Theory, and (6) speech intelligibility. Section 2 reviews some studies on NS conducted either worldwide or in Brazil. Finally, Section 3 wraps up the chapter.

### 2.1 LANGUAGE, COMMUNICATION, AND NUCLEAR STRESS

#### 2.1.1 Language

Language is not just an instrument for communication, but an environment with culture that creates us and shapes us. (Leffa, 2002, p. 11)

Defining language is as a complex task, as complex as language itself can be. In agreement with Larsen-Freeman (1997, 2007), the Chaos/Dynamic Systems theory seems to offer a way of looking at language which is deep and true to the complexity and dynamism of language.

Chaos/Dynamic systems are depicted according to some features, namely, dynamism, complexity, nonlinearity, chaos, unpredictability, sensitiveness to initial conditions, openness, self-organization, feedback sensitiveness, and adaptiveness, and to a variable called “strange attractors” (which are fractal) (Larsen-Freeman, 1997). Language can be considered a dynamic and complex system itself since much of these features, in different degrees, can be used to its depiction. The paragraphs that follow include a brief overview of some of these features as discussed by Larsen-Freeman (1997).

Language is dynamic mainly for three reasons: (1) its use involves an active process, (2) it grows/changes, and (3) it has a chaotic nature. Although “language can be described as an aggregation of static units or products”, the use of these units/products “in actual speech involves an active process, usually referred to as *parole* (Saussure) or performance (Chomsky)” (Larsen-Freeman, 1997, p. 147). Language is not static, but dynamic: it is compared to organisms given that it

develops and grows. This development and growth are use-dependent, that is, they depend on the natural selection of items that an individual makes during language use. Finally, it is dynamic because of its chaotic nature, which is related to the way language is typically construed: its current use and its changes triggered by use cannot be separate (i.e., they are isomorphic). Every time one writes/speaks and someone else reads/listens, changes occur<sup>6</sup>: “the rules are shaped by discourse. Thus, the behavior of the system as a whole is the result of the aggregate of local interactions” (Larsen-Freeman, 1997, p. 148).

These natural selections by language users (speakers), in conjunction with other factors (e.g., the advance of technology), trigger the language to develop/grow and thus change. These changes that language experiences diachronically are nonlinear, since “new forms enter and leave the language in a non-incremental fashion” (Larsen-Freeman, 1997, p. 147). They are also unpredictable: knowing when a new form enters or when an existing form leaves is somewhat impossible. Yet, they are not uniform, “so that a synchronic snapshot of language might appear chaotic” (Larsen-Freeman, 1997, p. 147): it may show many words meaning the same thing at a specific period in time (e.g., the case of reflexive pronouns). On the other hand, a diachronic look at the language may show the same word having distinct meanings (e.g., *awful*<sup>7</sup>).

Language is not only dynamic, but it is also complex, as it comprises many different subsystems, namely, phonology, morphology, lexicon, syntax, semantics, and, pragmatics, which are interdependent, that is, a change in one of them may “result in a change in the others” (Larsen-Freeman, 1997, p. 149). Looking into these interactions of the subsystems enables one to picture the whole of language (fractality).

Yet, language is dependent to initial conditions. Larson-Freeman (1997) explains that Universal Grammar is the initial condition of human language. Every language has some principles that constrain the shape of a given language. She illustrates this by drawing on voicing

---

<sup>6</sup> For instance, I have noticed that the verb “revolve around” has been used by some scholars (e.g., Larsen-Freeman, 1997; Gonçalves, 2014) and I have made the decision to use it in my writings, changing, to a certain extent, the way I collocate the words study, research, questions and the like.

<sup>7</sup> When it first appeared, the term *awful* was used to refer to a positive feeling (“awe-inspiring”). At some point in time, it started to be used to refer to something “extremely bad” and, now, it has been used to express something bad, but not to a great extent as before (Wijaya & Yeniterzi, 2011).

assimilation of obstruents, common in many languages, and on the assimilation of voiced consonants to the voiceless in English, which has a different behavior in Spanish and Russian. These principles are fields of attraction and their power differ in every language. Strange attractors are one of the many factors triggering language variation, which depends on the “strength each field [of attraction] exerts on a particular language” (Larsen-Freeman, 1997, p. 150).

Language is also fractal. This is well exemplified by Zipf’s power law in which word rank is connected with word frequency. According to Zipf’s power law, “if a word occupies a particular word frequency rank in a given language, then it is likely to reflect that same frequency *in any given text* of that language” (Larsen-Freeman, 1997, p. 150). In this sense, the words that appear in a portion of a text might appear in the other portions of the same text, being then fractal: “The pattern that exists at one level of scale holds for other levels and for the whole system” and the continuous change are also “reflected at every level of scale” (Larsen-Freeman, 1997, p. 150). Language is organic and live and, as Larsen-Freeman (1997) would say, as I am writing this chapter I am contributing to the changing frequencies of word use in this dissertation and in the whole system of English.

Given the complex and dynamic nature of language and language use, communication and interaction are, in turn, considered complex and dynamic, too. These concepts are addressed in Section 2.1.3, after English and speakers of English are defined (Section 2.1.2).

### **2.1.2 English language status and speakers of English**

EIL [English as an International Language] by definition no longer belongs to any one nation or culture. (McKay, 2003, p. 17)

The very fact that English is an international language means that no nation can have custody over it. (Widdowson, 1994, p. 385).

Globalization has influenced economy, technology, ideology, and the use of English language has mediated these areas. It has been used for supporting the movement of capital, goods, services, labor, and making possible the multifarious interactions among people around the world (Leffa, 2002). The use of English by the international community

has brought to question the ownership of the language (e.g., McKay, 2003; Widdowson, 1994) and the discussion of the status for the language. Each proposal reveals issues of sociopolitical power, hegemony, ideology, and geographical boundaries. The list is gigantic: *English as Native Language, as a Second Language, and as a Foreign Language* (Kachru, 1986, 1992; Kirkpatrick, 2007), *as a Medium of Intercultural Communication* (Meierkord, 1996, as cited in Seidlhofer, 2005), *as a Multinational Language* (Leffa, 2002), *as a Lingua Franca* (Firth, 1996; Jenkins, 2000; Seidlhofer, 2008), *as an International Language* (Jenkins, 2000; McKay, 2003; Sharifian, 2009), *as a Global Language* (Crystal, 2003), *World Englishes* (B. Kachru, Y. Kachru, & Nelson, 2006), *International English* (McKay, 2002), *new Englishes – nativized and indigenous varieties* (Kirkpatrick, 2009).

Even though the present study is aware of the historical, political, social and economic factors involving the choice of English for worldwide communication, which have heated up debates on the status of English, it chooses to simply refer to English as *International Language* (IL) as naively as possible, as it is not the scope of this work to tackle these issues<sup>8</sup>. My view of English as an International Language is an umbrella term that permeates three notions: (1) that of number (additional, second, third, and so forth in opposition to first language); (2) that of use (international – not favoring a specific English as an L1 variety<sup>9</sup>); (3) that of acquisition (interlanguage: it is shaped according to the L1 attractors and language use).

Therefore, the present study follows McKay's (2010) definition of English as an International Language whereby it is an "umbrella term to characterize the use of English between any two [IL] speakers of English, whether sharing the same culture or not, as well as between [IL] and L1 speakers of English" (McKay, 2010, p. 95). However, the

---

<sup>8</sup> However, I cannot prevent myself from stating my personal opinion on this. I agree with Bisong's (1995 as cited in McKay, 2003) view of English when talking about Nigerians learning English. I learn English not because I am forced to. I learn it because I want to hold interactions worldwide, to have access to the advances of technology, medicine, to have access to information in general and knowing English allows for the accomplishment of that. In no matter, English oppresses me or makes me feel inferior. Similarly, in no ways, knowing English makes me value less the L1 I speak. As Bisong would say, I am "sophisticated enough to know what is in" my interest, and that my "interest includes the ability to operate with two or more linguistic codes in a multilingual situation" (Bisong, 1995, p. 131 as cited in McKay, 2003, p. 5).

<sup>9</sup> English as an International Language "emphasizes that English, with its many varieties, is a language of international, and therefore *intercultural*, communication" (Sharifian, 2009, p. 2).

stand taken by the present study focuses on the definition of English as an International Language (EIL) as determined functionally by its use rather than by ownership.

Regarding the ownership of English, from my standpoint, the status *international* is not a matter of ownership (belonging to every community who speaks it), but that of usage (being used for international communication). It is undeniable that English is from a given country and that will never change. The fact that it is from that country, though, does not change the fact that the language can be used somewhere else in the world. Under the emergentist view of language (Section 2.1.1), it could be said that (1) English is of those who inherited<sup>10</sup> it from their ancestors and that, as any other language, (2) it changes every day. The changes are caused by its use by their L1 speakers as well as by their IL speakers, that is, “English is being shaped at least as much by its nonnative speakers as by its native speakers” (Seidlhofer, 2005, p. 339). It is not mine (but shaped by my use of it), it is not yours (but shaped by your use of it), it is not theirs (but shaped by their use of it), it is not ours (but shaped by every use we make of it).

In light of the aforementioned, in the present research, (1) Brazilian Portuguese participants will be referred to as L1 Brazilian Portuguese (L1-BP) speakers of English (they are from Brazil and they use – or aim at using – English as an additional language in international contexts), (2) participants who are speakers of English from other nationalities will be referred to as International Language (IL-English) speakers (they have their own L1, and English is an additional language of international use), and (3) speakers who have English as a first language will be referred to as first language English (L1-English) speakers of English (English is their L1, it does not matter whether they have any other additional languages). Note however that the L1-BP speakers of English are also IL-English speakers, and, together, BP, IL and L1 speakers of English are all speakers of English with different levels of proficiency. With regards to proficiency, every time it is important to make difference in proficiency levels, I will include low, intermediate, or high before BP or IL speakers of English and I will take

---

<sup>10</sup> How can I say that a language is not of someone else when my own language is not mine? It is inherited from my parents/family/community, who, in turn, received it as heritage of their ancestors. How can I say it is mine or it is not someone else's if every interaction that I hold with others or with a text (or with myself) that language is being changed?

for granted that L1-English speakers are highly proficient speakers of English as much as we, Brazilians, are highly proficient speakers of Brazilian Portuguese (BP)<sup>11</sup>. In addition, whenever defining a participant L1 background is necessary, the L1 will be placed before IL (e.g., French IL speakers of English).

Owing to the nature of language, as stated earlier in this chapter, and to the idiosyncrasy of each individual, communication is fragile. Having speakers from different L1 interact can enhance this fragility, yielding even more opportunities for communication breakdowns. The next portion of this chapter addresses issues of communication, interaction and reviews the Communication Accommodation Theory (CAT).

### **2.1.3 Communication, Interaction, and Communication Accommodation Theory**

The spirit is willing but the flesh is sometimes weak. (Jenkins, 2000, p. 177)

In a general sense, communication refers to the transmission and reception of INFORMATION (a ‘message’) between a source and a receiver using a signaling system: in linguistic contexts, source and receiver are interpreted in human terms [speaker/talker and listener/hearer, respectively], the system involved is a LANGUAGE, and the notion of response to (or acknowledgment of) the message becomes of crucial importance (Crystal, 2008, p. 89).

Speech communication is “a highly adaptive process on the parts of both the talker [speaker] and the listener” (Bradlow & Pisoni, 1999, p. 2075), that is, both parts (listener and speaker) make continuous adaptations to the demands of a given communicative situation. They develop strategies of communication (Seidlhofer, 2008) to negotiate

---

<sup>11</sup> I could refer to the participants as BPSE (Brazilian Portuguese speakers of English), ILSE (International Language speakers of English), and LISE (L1 speakers of English). However, because reading abbreviations are sometimes not reader friendly, I decided to use abbreviation for only Brazilian Portuguese (BP), International Language (IL), and first language (L1). For the sake of avoiding repetition, sometimes “English” in “speakers of English” will be omitted.



meaning, achieve cooperation and consensus, gauge each other's levels of linguistic and pragmatic competence, and adjust expectations on the linguistic and pragmatic levels. Communication Accommodation Theory (CAT) brings contributions to understanding the interactions held between speakers/listeners engaging the process of communication. The following paragraphs review some concepts of CAT.

Giles, N. Coupland, and J. Coupland (1991) trace the development of Accommodation Theory from (1) "a sociopsychological model of speech-style modifications" to (2) "an integrated, interdisciplinary statement of relational processes in communicative interaction (...), an "interface between language, communication, and social psychology" (p. 2). It has been synonyms with "listener adaptedness", "person-centered or other-related/directed speech", "taking the role/perspective of another", and "positive politeness" (Giles et. al., 1991, p. 2). The first tenets of Accommodation Theory were a response to insensitivity to social contextual variables in early sociolinguistic research.

One of the central variables for communicative success is language users' capacity for accommodation. According to Giles et al. (1991, p. 2), Accommodation "is a multiply organized and contextually complex set of alternatives, ubiquitously available to communicators in face-to-face talk". The scholars also highlight that accommodative processes can facilitate or impede proficiency in an AL. There are two main strategies involved in the accommodation process, namely convergence and divergence.

Convergence refers to a strategy in which language users adjust to "each other's communicative behaviors in terms of a wide range of linguistic-prosodic-nonverbal features including speech rate, pausal phenomena and utterance length, phonological variants, smiling, gaze, and so on" (Giles et al., 1991, p. 7). To illustrate, let us take as an example a simulated IL interaction between two interlocutors, namely, an L1-English and an IL-English speaker, whereby during the many turns taken by both, the L1-English speaker notices, after some interruptions for negotiation of meaning, that the process of flapping<sup>12</sup> undertaken by him/her is causing breakdowns in communication when he/she pronounces words such as *water*, *butter*, *lettuce*, and *cheddar*.

---

<sup>12</sup> The term *flapping* (or tapping) refers to a process that converts an alveolar stop (e.g., /t/) to a voiced flap (or tap) (/r/) (Yavas, 2011).

The L1-English speaker then decides to avoid this process and keeps the /t/ or /d/ pronunciation in the words to enhance his intelligibility to that IL-English speaker. By doing so, the L1-English individual as a *speaker* is converging with the IL-English individual as a *listener*.

Divergence, in turn, refers to a strategy in which language users “accentuate speech and nonverbal differences between themselves and others (Giles et. al., 1991, p. 8)”. In our simulated interaction between the L1-English and the IL-English speakers, if instead of avoiding the use of flapping, the L1-English speaker emphasized its use in order to, for instance, pose as superior, he would be then diverging from the IL-English individual as a listener. In this sense, communication, if not broken, would be hard and unpleasant. Divergence is thus seen as a dissociative communicative act.

In international communication, it is more common to believe that convergence, or at least the desire to converge, will be more found as interlocutors’ aim is the same: to reach success in conveying their messages – to understand and be understood. However, when one problematizes this issue, three aspects should be considered. The first one refers to the type of interaction being held, the second and the third refer to the lack of resources available to accommodate.

When one thinks of interaction, in linguistic terms, what may come to ones’ mind is the image of two or more people involved in a speech act. According to Crystal (2008, p. 248), interaction refers to the “study of speech in face-to-face communication”. Communication, however, is not a merit of only more than two people present at the speech act on face-to-face communication. Communication of messages also happens in situations other than face-to-face (e.g., informative listening), such as through watching TV, listening to the radio, and watching videos in the Internet, where the speaker cannot see the reactions of the listener, who, in turn, is not able to negotiate meaning with the speaker. So, even though the speaker may think of accommodating, what shape the accommodation might take is not precise and impossible to estimate.

The second and third issues have to do with the resources available, true mainly for interactions held between IL-English speakers. In such interactions, accommodating to one’s interlocutor’s speech is a way to guarantee intelligibility. However, when something may not appear to make sense to IL-English speakers, “they are unable to decide whether the speaker has used a word with which they are unfamiliar,

whether they have misheard, or whether the speaker has, in fact, made an error” (Jenkins, 2000, p. 82) and thus may not take a step into solving the problem through negotiation and accommodation (issue 2). Additionally, even though an IL-English speaker can be motivated to accommodate, he/she may lack the resources to do so (issue 3), that is, the speaker may lack the phonological knowledge which is important to his or her intelligible discourse (Jenkins, 2000; Zuengler, 1991). Due to this lack of resources, research on intelligibility has attempted to map the pronunciation features that are a must to be respected in order to guarantee an IL-English speaker’s intelligible speech. Once these features are understood and considering that “the sounds that are likely to undergo accommodative shifts may be those that the speakers are most conscious of” (Zuengler, 1991, p. 233), pedagogical practices that bring these sounds up to discussion with pronunciation variations awareness raising activities should promote learners with resources on pronunciation to be used when interacting in international contexts (Zuengler, 1991). The next portion of this section discusses communication in terms of success and breakdowns and the way intelligibility is understood in this research.

#### **2.1.4 Speech Communication, Perception, and Intelligibility**

We need to continually modify and fine-tune our language in order to communicate with other people. (Seidlhofer, 2009, p. 196)

When one is able to fine-tune one’s language in order to communicate well with one’s interlocutor and one’s interlocutor, in turn, fine-tune his/her own language in order to understand his/her interlocutor’s intended message, the two are bound to reach successful communication. Successful communication takes place when listeners respond to a speaker’s utterance “in a manner which is appropriate to his [the speaker’s] purpose in speaking” (Catford, 1950, p. 7). As Catford (1950) sees it, this appropriateness is called “effectiveness” and differs from “intelligibility”.

Speech is generally said to be *intelligible* if the hearer understands the words, i.e. if his response is appropriate to the *linguistic forms* of the utterance: that is to say, if it is in accordance with

the semantic habits of the speech-community whose language is being used. An utterance may be *intelligible* in this sense, yet *ineffective* in the sense that the hearer's response is not what the speaker intended. (Catford, 1950, p. 8)

An intelligible utterance can then be ineffective, but the other way round is not possible.

Table 1  
*Combinations for speaker's utterance intelligibility and effectiveness*

TYPE OF UTTERANCE	HEARER'S RESPONSE
(1) Unintelligible+Ineffective	None, or inappropriate
(2) Intelligible+Ineffective	Inappropriate
(3) Unintelligible+Effective	(Appropriate, only by chance)
(4) Intelligible+Effective	Appropriate

(Source: Catford, 1950, p. 8)

In accordance with Catford (1950), the present research sees that successful communication takes place if the speakers' utterances are both intelligible and effective (combination 4), which leads to appropriate responses from the part of the listeners. In other words, the speaker needs to use appropriate linguistic forms (community and situation dependent), that is, words, morphological and syntactical devices, sounds and prosody, in an appropriate manner in order to attain intelligibility and effectiveness.

Communication is unsuccessful, in turn, when there are communication breakdowns; in other words, when speech is unintelligible and ineffective (combinations 1 and 2). This loss of intelligibility leading to unsuccessful communication may be a function of the following factors:

- (1) Speaker's *selection* of linguistic forms. With regard to pronunciation, if the speaker selects, for example, a sound which does not match the sound that should be pronounced, it may cause him/her to produce it in a different fashion. For example, when saying *Do you like rock?* instead of choosing the English /r/, he/she selects one of the BP sounds for the grapheme <r>, that is, the /h/, saying then *Do you like hawk?*. The listener, having *hawk* available in his/her lexicon, will certainly think of the kind of bird

rather than the kind of music, which will lead him/her to provide an inappropriate response, and have a communication breakdown.

- (2) Speaker's *execution* of linguistic forms. Sometimes, the speaker may choose a sound which, in his phonological system, seems to be adequate. To illustrate, let us take the word *tea*. The speaker makes the correct selection of /t/. However, when executing it, he/she does not succeed to avoid affrication, pronouncing [tʃi:]. The listener can then understand *Would you like some Chee?* which could be understood as Cheetos®, cheese, cheesecake if these were also available at the table (Catford, 1950). Prosodically, speakers that tend to produce some pitch change in utterance initial position as a trait of their speech, when having to signal information by means of NS in other utterances position may have their intents misinterpreted. For example, if they answer the question *What did Maria do with the message?* by saying *Maria deleted the message* with some pitch change on *Maria* and another pitch change in *deleted*, speakers would have their intents misinterpreted. Their listeners might understand that the speakers want to imply that Maria deleted the message while some other people might have forwarded it.
- (3) *Transmission* from speaker to hearer. Supposing the speaker selected and executed the utterances in an expected manner, but the environment was too noisy, it is possible that because of that noise communication could be broken.
- (4) Hearer's *identification* of linguistic forms. The use of flapping by the speaker with a listener who does not have that process in his phonological system will cause him/her not to understand the speaker's utterances, leading to communication breakdowns. Prosodically, the use of NS in utterance initial position and another signaling important information in non-initial utterance position with a listener who does not have initial-pitch-changing as a trait of his/her speech may cause the listener to misjudge the speaker's intent.
- (5) Hearer's *interpretation* of linguistic forms. Having not identified a given sound, the listener can interpret the message uttered as anything else, as explained in action (2). The same would happen prosodically if the listener is unable to identify the pitch change signaling important/new information, as also explained in action (2).

These points interplay and overlap (1 and 2, 4 and 5). That is the reason why the more one's phonological system is complete, the more resources and situations for practice an individual may have, the more chances he/she will have to accommodate and to avoid breakdowns in communication, if willing to.

Provided that IL-English speakers have unlimited access to interaction, which decades ago was granted mainly when visiting an English-speaking country (Catford, 1950), they will no longer play the sole role of a listener. The advances of technology brought in chatting tools such as *Facebook*®, *Skype*®, and *Whatsapp*®, which have definitely allowed IL-English speakers to play both roles, that of speaker and that of listener, more often. Because of that, developing both speaking and listening skills is definite to accommodating, being intelligible and finding someone else intelligible.

Catford (1950) was one of the first to define intelligibility (and effectiveness), and in the past six decades, many other definitions appeared, to mean sometimes the same thing at the same level or at different levels: *intelligibility* (Catford, 1950; Smith & Rafiqzad, 1979; Smith & Nelson, 1985; Kenworthy, 1987; Bamgbose, 1998; James, 1998; Jenkins, 2000; Field, 2003); *effectiveness* (Catford, 1950); *comprehension* (Smith & Rafiqzad, 1979); *comprehensibility* (James, 1998; Smith & Nelson, 1985; Field, 2003; Derwing & Munro, 1997), *interpretability* (Smith & Nelson, 1985), *understandability* (Kenworthy, 1987), *communication* (Kenworthy, 1987), *accessibility* (Dalton & Seidlhofer, 1995), *acceptability* (Dalton & Seidlhofer, 1995), *communicativity* (James, 1998). Some of the terms seem to be interchangeable as they refer to the same construct while some do not. To illustrate, let us take *interpretability* (by Smith & Nelson, 1985), *communication* (by Kenworthy, 1987), and *communicativity* (by James, 1998), which resemble *effectiveness* by Catford (1950), that is, these terms refer to the meaning behind the word or utterance and is deprehended by the listener's response. Conversely, while *comprehension* by Derwing and Munro (1997) refers to how difficult/easy it is to understand an utterance/a word, for Smith and Nelson (1985) it refers to understanding the meaning of the utterance/word.

Perhaps the most comprehensive definition was that proposed by Smith and Nelson (1985), breaking down the term into three different dimensions: intelligibility, comprehensibility, and interpretability. These

dimensions refer to three levels, respectively: (1) word and utterance RECOGNITION (PHONOLOGICAL LEVEL); (2) word and utterance MEANING (SEMANTIC LEVEL); and (3) grasp of speaker's INTENTIONS (PRAGMATIC LEVEL). Bamgbose (1998) acknowledges the three levels in his definition for intelligibility: "a complex of factors comprising recognizing an expression, knowing its meaning, and knowing what that meaning signifies in the sociocultural context" (Bamgbose, 1998, p. 11). In the present study the three levels are also considered and a choice was made for using the term *interpretability* to refer to the result of recognizing the utterance, its meaning and the consequent speaker's intent grasp.

Based on the concepts mentioned in this section, in the present study, successful communication refers to communication acts in which the speakers' utterances are both recognizable, meaningful and interpretable according to the listener's view (effective in Catford's terms), while communication is considered unsuccessful when the speaker fails to utter interpretable utterances that yield inappropriate responses from the part of the listeners. In the present study, unsuccessful communication was classified into two different levels: (1) sights of reduced interpretability and (2) real communication breakdowns. Utterances with sights of reduced interpretability are those delivered with some noise (unexpected pauses, pronunciation, syllable lengthening, intonation, NS placements), that affected the listener's perception of nuclear stress (defined in the next section) but, which after some effort to negotiate meaning by means of repetition, had the noise lowered down and the expected message understood. Real communication breakdowns, in turn, refer to the unexpected interpretations of the speaker's intent, with or without meaning negotiation.

### 2.1.5 Nuclear Stress

Accent [nuclear stress] is predictable (if you are a mind reader). (Bolinger, 1972, p. 633)

Terminology involving prosody may at times be somewhat confusing. At this point, it is important that a distinction between word stress, sentence stress, nuclear stress, and intonation be drawn.

In the present study, *word stress* simply refers to the most prominent syllable in a word. That most prominent syllable can be usually established a priori in word citation form found in the dictionary<sup>13</sup> (Celce-Murcia, Brinton, & Goodwin, 1996).

*Sentence stress* is the prominence placed in content words in a sentence or utterance. It has a grammatical basis: Content words are more prominent than function words. Sentence stress is responsible for the rhythm of the language (Celce-Murcia et al, 1996). In the literature, *pitch accent* is a term often used to refer to both *sentence stress* (e.g., Pierrehumbert, 1980) and *nuclear stress* (e.g., Gussenhoven, 1994) and that is sometimes confusing. For this reason, the present study made the choice to stick to *sentence stress*.

The concept of *thought groups* is important to the comprehension of nuclear stress. It has been used in the literature under different names, such as *Intonation Units* (Hirst, 1998), *Intonational Phrases* (Frota & Moraes, 2016; Heusinger, 2007; O'Mahony, 2014; Pierrehumbert, 1980), and *tone units, sense groups, tone groups, and breath groups*, as listed by Pierrehumbert. In the present study, the choice for *thought groups* was made because of the clear match between the lexicon comprising the term and its meaning: Thought groups are meaningful parts of speech set off by pauses before and after, with a nuclear stress, an intonation contour of their own and a coherent internal structure (Celce-Murcia et. al., 1996). In order to a thought group to be formed, some phonetic, syntactic, and semantic criteria exist, as listed by Heusinger (2007, p. 280), transcribed below.

Linguistic criteria for defining an intonational phrase (IP) [thought group]

- (i) *Timing*: An IP [thought group] can be preceded and followed by a pause.
- (ii) *Metrical*: The metrical structure provides an additional clue, viz., the presence of a most prominent accent [NS].
- (iii) *Tonal*: The boundary of an IP [a thought group] is sometimes tonally marked by a boundary tone. Pitch range adjustment plays a role, as well.
- (iv) *Junctural*: The boundary of an IP [a thought group] can block certain junctural phenomena (cf. Nespor & Vogel (1986)) (*sic*).

---

<sup>13</sup> Word stress can be predicted, but it varies according to a range of factors, such as English variety. For an extensive description of word stress rules, refer to Roach (2009).



- (v) *Syntactic-prosodic*: The boundaries of an IP correspond to those of some syntactic constituents.
- (vi) *Semantic*: The material in the IP must constitute an *informational unit* or *sense unit*.

According to Heusinger (2007), these criteria can sometimes conflict and the author highlights that from these the most reliable criterion is the boundary tone (L% or H%<sup>14</sup>). Languages such as German that not always show a boundary tone, can make use of “pauses, lengthening of the final syllable and a pitch accent for each intermediate phrase” (p. 280). The scholar argues that the discourse function is one of the most reliable criteria.

The length of a thought group may be short or long, depending on its appropriateness for establishing a coherent discourse (Heusinger, 2007). Despite this being so, according to Crystal’s (1969, as reviewed by Hirst, 1998) description of his corpus, the average length of thought groups was of five words and that 80% of the thought groups had at maximum eight words: Longer utterances are usually broken up into two or more thought groups (Hirst, 1998). Having this in mind, utterances in the present study had at maximum six words (if function words are also computed).

*Nuclear stress* itself has also received many names: *Nuclear or tonic stress* (Jenkins, 2000), *nuclear stress* (Pierrehumbert, 1980; Zubizarreta, 2016), *prominence* (Celce-Murcia et al, 1996), *phrasal stress* (Pierrehumbert & Hirshberg, 1992), *accent* (Bolinger, 1972, 1998; Frota & Moraes, 2016; Sluijter & van Heuven, 1996a), *sentence stress* (Galaczi, Post, Barker, & Schmidt, 2017; Matic, 2015), *contrastive stress* (Jenkins, 1997), *pitch accent* (Gordon, 2014; Gussenhoven, 1994; 2007; Krahmer & Swerts, 2007), and *nuclear accent* (Mateus, 2004; Ortiz-Lira, 1998). Such varied terminology is used to refer to the placement of more prominence on a specific syllable/word/phrase in a thought group, in order to highlight it and convey (1) meaning, (2) the context in which the utterance is placed, and (3) the speaker’s intent (Celce-Murcia et al, 1996). Nuclear stress, thus, has great communicative value. While sentence stress is grammar oriented (content VS function words), nuclear stress is context oriented,

---

<sup>14</sup> L% refers to a low boundary tone while H% refers to a high boundary tone. They are used to mark the end of a thought group with a falling pitch movement in the case of the former and with a rising movement in the case of the latter (Pierrehumbert, 1980).

that is, its allocation depends on the intentions of the speaker. Sentence stress and nuclear stress sometimes coincide, but it is not always the case (Lambrecht, 1998; Zubizarreta, 2016). To illustrate let us consider the sentence in (1).

(1) She ASKED me to STAY<sup>15</sup>.

When said in an unmarked context, that is, as neutral information, the nuclear stress coincides with the rightmost sentence stress in (1), namely, STAY. However, if the same sentence is uttered in an attempt to emphasize who the person being asked to stay is, a different realization of nuclear stress is then made, as illustrated in (2).

(2) She ASKED ME to STAY.

Note that in (2) greater prominence is placed on ME, a word that in (1) takes no sentence stress, meaning that a female asked the speaker to stay rather than her brother, for instance. This perceived prominence of the word holding the nuclear stress (NS) is manifested by means of changes in pitch (Sluijter & van Heuven, 1996a) “usually combined with increases in length and loudness<sup>16</sup>, imposed, as a rule, on the ‘stressed’ syllable of a word that is to be highlighted” (Bolinger, 1998, p. 47). *Pitch* refers to the relative volume (altitude) of the sound (Nooteboom, 1997), that is, how low or how high one perceives the sound during the speech flow. The physical correlate of pitch is the *fundamental frequency* (F0), which is “determined by the rate of vibration of the vocal chords” (Mary, 2012, p. 12). The range of this rate varies among individuals. In conversational situations, while for males it usually goes from 80 to 200 Hz, for females the range is usually from 180 to 400 Hz (Nooteboom, 1997). The direction of the F0 change can be falling or rising. Pitch can be tracked in specific software such as Praat (Boersma & Weenink, 2014) through the capture of the F0 in voiced segments.

---

<sup>15</sup> Conventions: lower case letters = unstressed syllables; SMALL CAPITAL LETTERS = stressed syllable with sentence stress; LARGE CAPITAL LETTERS = nuclear stress.

<sup>16</sup> Loudness is one of the response functions of our hearing mechanisms and it refers to the intensity of the sensation from a sound or noise. This sensation can vary in scale from soft to loud. Loudness is not affected only by sound pressure, but also by the frequency and the waveform of the sound (Olson, 1972). Its physical correlate is amplitude.

*Intonation*, in turn, refers to the controlled modulation of pitch, a sequence of sound units that vary among low and high tones<sup>17</sup> and thus gives some distinguishable melodic properties to speech (Mary, 2012; Ortiz-Lira, 1998).

A NS can be produced with different tunes<sup>18</sup> (falling-rising F0 or rising-falling F0) and still be the NS (Pierrehumbert & Hirshberg, 1992). That means that a given tune (rising or falling) does not make nuclear stresses distinctive from one another. The nuclear stress is where the pitch contour<sup>19</sup> usually begins, as illustrated in Figures 1, 2, 3, and 4, reprinted from Pierrehumbert (1980) by Pierrehumbert & Hirshberg (1992).

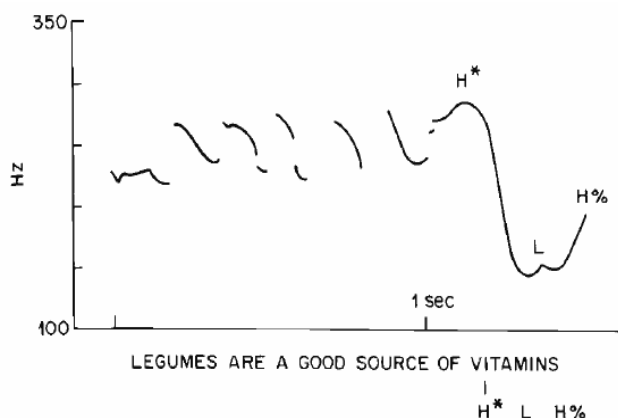


Figure 1: *Legumes are a good source of VITAMINS*. Production with the NS on vitamins, with a falling-rising intonational pattern (Pierrehumbert & Hirshberg, 1992, p. 273).

In Figure 1, the NS is in the word *vitamins*. Note that the pitch movement starts in the stressed syllable ‘VI’ in *vitamins*, signaling the

<sup>17</sup> A tone is the distinctive pitch level of a syllable (Crystal, 2008).

<sup>18</sup> Tunes refer to structured strings of low or high tones (Pierrehumbert, 1980).

<sup>19</sup> Pitch contour is the fundamental frequency curve (prompted by pitch changes) as the ones displayed in Figures 1 to 4 in the present section.

presence of the NS. The intonation pattern for the pitch curve is H\*LH%, that is, the stressed syllable holds a high tone (H\*), which is followed by a low tone in the next unstressed syllable (L) and finally by a high boundary tone in the last unstressed syllable (H%).

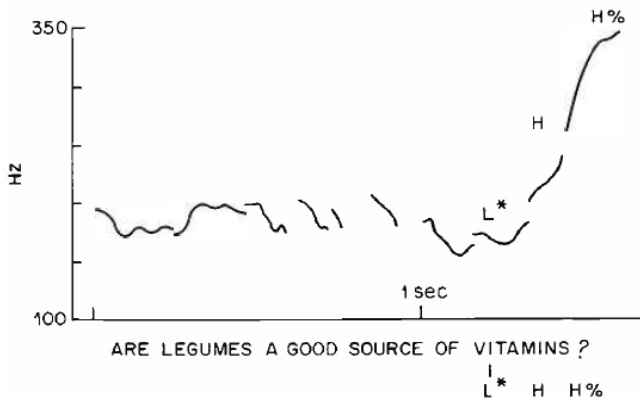


Figure 2: *Are legumes a good source of VITAMINS?* Production with the NS on vitamins, with a rising intonational pattern (Pierrehumbert & Hirshberg, 1992, p. 273).

Note now that again, for this question in Figure 2, the NS is in the same position as in Figure 1, namely on the word *vitamins*. The tune, however, is different (L\*HH%). It now starts with a low tone in the stressed syllable 'VI' (L\*), which also signals the NS, followed by a high tone in the next unstressed syllable (H) and then a final boundary tone in the last unstressed syllable (H%).

When NS is allocated in order to highlight another piece of information, such as that legumes are a GOOD and not a BAD source of vitamins, the pitch contour starts up earlier in the sentence as shown in Figure 3.

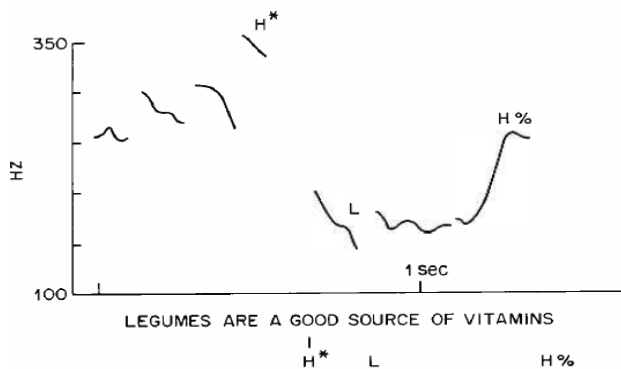


Figure 3: *Legumes are a GOOD source of vitamins*. Production with the nuclear stress on GOOD with a falling-rising pattern (Pierrehumbert & Hirshberg, 1992, p. 274).

Note that in Figure 3 the pitch contour starts in *good* and finishes in *vitamins*. However, the NS is only on the word GOOD. The intonational pattern this time is H\*LH%, the same found in Figure 1, but it is distributed along the phrase *GOOD source of vitamins*. Figure 4 shows the same position for NS in a polar question<sup>20</sup>.

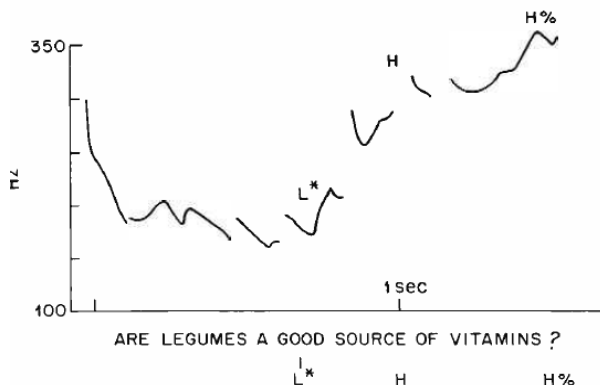


Figure 4: *Are legumes a GOOD source of vitamins?* Production with the NS on GOOD with a rising pattern (Pierrehumbert & Hirshberg, 1992, p. 273).

<sup>20</sup> Polar questions are often referred to as *yes-no questions* in textbooks.

Observe again that, in Figure 4, the pitch movement starts on the word *good* and finishes on *vitamins*, as it does in Figure 3, but now the NS has a low tone (on *GOOD*), followed by a high tone on the word *source* and final high boundary tone on *vitamins* (L\*HH%). It shows the same intonational pattern displayed in Figure 2, but with a different distribution<sup>21</sup>. These examples show that although a NS has a tone associated to it, its placement is not tone dependent, that is, pitch change range is what matters for NS, not the intonational pattern that it shows. The next section explores nuclear stress placement in more detail.

Although some reference is made to intonation, mainly in relation to the production of questions, the present study is concerned with NS allocation and does not set out to describe the tunes in pre nuclear and nuclear position, that is, it is not concerned with the intonation. Despite the fact that word stress and the pronunciation of segments are not the focus of the present study, some comments about word stress and pronunciation are added in the discussion of the results in order to explain some of the unexpected productions in the dataset.

### 2.1.5.1 The function of nuclear stress

NS is used to highlight important information, which is usually information that is ignored by the listener, so it is generally considered new information, as exemplified in (3).

(3) A: What time does your bus leave?

B: It leaves at TEN P.M.<sub>[NS]</sub>

In the answer in (3), the constituent *at TEN P.M.* represents new information, provided that the remaining portion of the sentence *It and leaves* were also part of the question. Even though *it* was not in the question, it does refer to a constituent contained in the question, namely, *your bus*. These constituents *it* and *leaves* are then considered to be present in the listener's mind, and thus they cannot be considered new. Therefore, they are words that convey old or given information and should take no focus by means of NS.

The dichotomy old/new has been discussed in the literature. Lambrecht (1998), for instance, posits that the illustration in (3) is a

---

<sup>21</sup> For more detailed information on the elements of an intonational pattern (pitch accent, phrase accent, and boundary tones), refer to Pierrehumbert (1980).

simplistic way of looking at the dichotomy of new/old information. For the scholar, the status of the information being “conveyed by a proposition cannot be factored out and matched with individual sentence constituents” (p. 49), in the case of (3), *it leaves* (supposedly old) and *at TEN P.M.* (supposedly new), but rather with the “establishment of ... pragmatic relations that makes information possible” (p. 49). As the author would see it, the information conveyed by the answer in (3) would not be *at TEN P.M.*, but rather “the time that the bus I am taking leaves is ten P.M.”. The scholar argues that such an answer as the one in (3) could not function as an interpretable answer without the associated full proposition. Both *it leaves* and *at TEN P.M.* is information shared by both speaker and listener. What makes the latter new information is the relation it has with the former. The new information (the assertion) establishes a time relation between the two pragmatically presupposed propositions (that the listener is taking a train; that the train leaves at a certain time).

My intention here is not to review and assess the extensive literature on this matter. Rather, it is to state how the notion is to be understood in the present study, which takes a different stand from that of Lambrecht’s (1998). The present study simply relates the term *new* to the *ignorance* of a fact. It is not *new* in the sense that the addressee does not have it as a possibility in his mind, as claimed by Lambrecht, but rather as the information missing in a given pragmatic context. The set of examples (Lambrecht, 1998, p. 260) below show cases of both my view and his view.

(4) Q: Who did Felix praise?

A: a. He praised HIMSELF.

b. He praised YOU.

c. He praised his BROTHER.

d. He praised a woman you don’t KNOW.

e. He praised NOONE. (*sic*)

In (4a), (4b), and in (4c), the NS constituents activate a referent in the addressee’s mind, while in (4d) and (4e) the activation does not happen, because the addressee does not know the woman and no one was praised. Lambrecht (1998) argues that, only in the cases of (4d) and (4e), newness exists, because no referent activation is made possible. However, as I see it, from (4a) to (4e), new information is present, since

the information conveyed was not available to the addressee. It is not new in the sense that the addressee does not have it as a possibility in his mind, but rather as the information missing in that pragmatic context<sup>22</sup>. Along the study, *new* will be the term used to refer to this kind of information.

NS use greatly depends on the context and the intention of the speakers. Therefore, any syllable in the thought group can be nuclear-stressed in order to express focus by placing a pitch movement on it (Sluijter & van Heuven, 1995). *Focus*<sup>23</sup> is a term used to refer to non-presupposed material in a sentence which is of interlocutors' communicative interest (Crystal, 2008) and it may be manifested prosodically (by means of NS) and/or syntactically (by means of the use of cleft sentences, for instance) or morphologically (by means of the addition of a word marker that indicates focalization) (Elordieta, 2007; Gussenhoven, 1994). The present study investigates NS to convey focus prosodically<sup>24</sup>.

According to the organization of information, focus can be *narrow* (marked) or *wide/broad* (unmarked) (Ortiz-Lira, 1998;

---

<sup>22</sup> In his own way, Lambrecht (1998) kind of agrees with my stand when he posits that “the communicative purpose of the various replies in (5.28) [4] is to identify for the addressee the referent of a missing argument in a pragmatically presupposed proposition (...). And the function of the focus accent is to mark a particular constituent as the one designating that referent. Whether the referent of the missing argument is “new” or “old” in the discourse is irrelevant from the point of view of the prosodic form of the sentence” (p. 261). Again, he is only being too strict with respect to the terms new or old.

<sup>23</sup> Mello & Silva (2015) go further to make a distinction between *focus* and *emphasis*. For them, *focus* only signals contrastiveness and exhaustiveness. *Emphasis* would then apply to the other cases, such as reinforcement, that syntactically is made by adjectives (e.g., *baita*) and prosodically made just as focus. The example they provide is “Eu não FUI não” ‘*Me no I did not GO*’. The present study, however, views that there is no clear cut between focus and emphasis. In the context presented, the speaker could have wanted to emphasize s/he did not go in contrast with the idea that s/he would go. Therefore, it is more related to points of view. The present study would rather think of emphasis to signal important information with the aim to correct, elicit and contrast, irrespective of being formally called focus or emphasis, comment or topic, or any other terminology found in the literature for this phenomenon.

<sup>24</sup> While Lambrecht (1998) works with three categories for information structure, namely *argument focus*, *predicate focus*, and *sentence focus*, I work with focus on sentence positions, namely, initial (usually on the subject), medial (usually on the verb), and final (usually on the object). It is not the focus of this study to scrutinize the information structure of utterances, but rather understand the way information is asserted and comprehended by means of NS placement.



Zubizarreta, 2016). In wide-focus contexts, the utterance contains all-new information, such as when answering the question in (5).

(5) A: What's the matter?

B: JOHN<sub>[sentence stress]</sub> has MOVED<sub>[sentence stress]</sub> to CANADA<sub>[NS]</sub>.

The NS in the answer in (5) falls in the rightmost sentence stress available (*Canada*), given that all information provided here is regarded as new. Such a question was probably triggered by the sadness in the interlocutors' eyes, which is explained by the fact that John has moved to Canada, a piece of information ignored by Speaker A as being the reason for Speaker B's sadness. Narrow-focus utterances, in turn, contain both given and new information, as illustrated in (6).

(6) A: Who has moved to Canada?

B: JOHN<sub>[NS]</sub> has MOVED<sub>[sentence stress]</sub> to CANADA<sub>[sentence stress]</sub>.

In (6), 'John' is new information while 'has moved to Canada' is given information (available in the question) (Ortiz-Lira, 1998). In this context, Speaker A has probably heard that someone moved to Canada and s/he aims to know who the person is. In the answer, the focus is narrowed down to 'John' only, the information being elicited.

Based on meaning and form, Gussenhoven (2007) mentions eight types of focus: *presentational*, *corrective*, *counterpresuppositional*, *definitional*, *eventive*, *contingency*, *reactivating*, and *identificational* focus. I discuss below these concepts and make a match between this terminology and the three uses of NS employed in the present study, namely, for information being *elicited*, *corrected*, and *contrasted*.

*Presentational focus* refers to the general concept of focus. In it "the focus constituent is the part of the sentence that corresponds to the answer to a question, either overt or implied" (Gussenhoven, 2007, p. 91). In (7), Speaker B overtly answers the inquiry, that is, when the singer Adele would be in Canada. Therefore, NS is placed on *in December* and has the function to elicit that information. In the present study, NS used with this purpose is referred to as NS used to signal information being *elicited*.

(7) A: When will Adele be in Canada?

B: Adele will be in Canada IN DECEMBER<sub>[NS for eliciting information]</sub>.

*Corrective focus* highlights a “constituent that is a direct rejection of an alternative, either spoken by the speaker himself (...) or by the hearer” (Gussenhoven, 2007, p. 91). In (8), Speaker B is directly correcting Speaker A in relation to the item that John purchased (a BIKE, not a car). Therefore, NS is assigned to the word *bike* and has the function to correct the previous mistaken information provided by Speaker A. NS used with the purpose of correcting mistaken information is referred to as NS used to signal information being *corrected* in the present study.

(8) A: John bought a car.

B: No, John bought a BIKE<sub>[NS for correcting information]</sub>.

*Counterpresuppositional focus* highlights a constituent that has an indirect correction of a piece of information said/heard, as illustrated in (9), transcribed from Gussenhoven (2007, p. 92), first used by Ladd (1980).

(9) a. (A: Has John read *Slaughterhouse Five*?)

B: John does [n't]<sub>counterpresup</sub> READ books.

b. (A: I'm telling you: John reads books!)

B: I'm sorry, John does [NOT]<sub>corrective</sub> /DOES[n't]<sub>corrective</sub> read books.

In (9b), Speaker B presents a direct correction for “John reads books”, while in (9a) Speaker B works with the presupposition that John has not read the *Slaughterhouse Five* because John does not read books. In the present study, the use of NS for correcting information only directly (corrective focus) is investigated.

*Definitional focus* highlights information that “does not refer to a change in the world, but informs the hearer of attendant circumstances” (Gussenhoven, 2007, p. 92) and it requires stress on both subject and predicate. In (10b), there is a definitional focus, since it describes and highlights a permanent fact (that explains, for example, the reason why

the speaker loves the listener) while in (10a) it is *eventive*, that is, it highlights and describes a momentary state (the redness in the listener's eye) (both examples are from Gussenhoven, 2007).

- (10) a. [Your EYES are red]<sub>EVENTIVE</sub>  
 b. [Your EYES are BLUE]<sub>DEFINITIONAL</sub>

*Contingency focus* resembles eventive focus. Both provide information, but in the former the information is presented as *potentially relevant*" (Gussenhoven, 2007, p. 94). In (11a) reprinted from Gussenhoven (p. 94), it is implied that if there are dogs, they must be carried and focus is then contingent, with NS placed in *dogs* and *carried*. In (11b), it is implied that the speaker might be concerned because s/he does not have a dog (so, s/he might not be granted access), as explained by Gussenhoven. NS is placed only on *dogs*.

- (11) a. [DOGS must be CARRIED]<sub>CONTINGENCY</sub>  
 b. [DOGS must be carried]<sub>EVENTIVE</sub>

*Reactivating focus* refers to the highlight of previous mentioned information (old/given information). In (12), there is also the process of topicalization. NS is placed on both *John* and *dislikes*.

- (12) (A: Does she know JOHN?)  
 B: JOHN she DISLIKES.

*Identificational focus* is used to identify a constituent, which is carried out by using clefting combined with prosody. In (13), there is an illustration of this type of focus (Gussenhoven, 2007, p. 96). In order to provide information about whom the female dislikes, Speaker B uses the cleft structure *It is...* and the name of the disliked person *John* with the NS.

- (13) [A: I wonder who she dislikes]  
 B: It is JOHN she dislikes

Only two out of the eight types of focus explained above have a match with the NS uses in the present study, namely, *presentational*, NS used for information being *elicited*, and *corrective*, NS used for

information being *corrected*. The third use of NS investigated here is for information being *contrasted*. Jenkins (1997) uses *contrastive* to refer to NS in a position other than the canonical. Moraes (1998) and Frota and Moraes (2016) use the term *emphasis for contrast* (1998) or *contrastive focus* (2016) to refer to the use of NS to both correct and contrast information probably based on the reasoning that correction implies contrast in some manner. Hedberg and Sosa (2007), by drawing on the differences in the combination of tones (low or high), discuss the (im)possible distinction between *contrastive focus* and *contrastive topic*, being *topic* and *focus* two concepts with which information structure theory is concerned. In the present study I do not mean to specify whether nuclear stress is associated with the information structures of topic and focus or whether a specific pitch contour marks topic or focus as discussed by Hedberg and Sosa (2007).

The present study preferred to treat *correction* and *contrast* separately, according to their functions, irrespective of information structure. In the NS placed for *correcting* information, there is a mistaken piece of information in the discursive context that needs correction. However, in utterances with NS placed for *contrasting* information, there is no mistaken information, but possibilities, as illustrated in (14).

(14) (Context: Discerning on a person to be invited to go parachuting.)

Speaker: Is ZENNY<sub>[NS for contrast 1]</sub> afraid of heights or is it ANNA<sub>[NS for contrast 2]</sub>?

The question in (14) contains two clauses connected by a disjunction. NS is placed both on *Zenny* and on *Anna* to show the contrast, that is, to signal that the speaker is in doubt about the person who is afraid of heights, an important piece of information to enlighten the decision of whom to invite. The next section presents how NS is manifested in English and in Brazilian Portuguese in relation to the structures containing the three NS uses investigated.

### 2.1.5.2 Nuclear Stress in English and in Brazilian Portuguese

As previously mentioned in Section 2.1.5, NS is manifested by means of changes in pitch usually in conjunction with syllable

lengthening and increased loudness of the constituent on which the NS falls (Bolinger, 1998; Sluijter & van Heuven, 1996a; Klok, Wagner, & Goad, 2011) in English. This manifestation of NS is shared by BP language (Carpes, 2014; Frota & Moraes, 2016; Lucente, 2012; Mello & Silva, 2015).

The portion of the sentence whereby the nuclear stress falls greatly depends on the context and the intention of the speakers (Zubizarreta, 2016). In general, it occurs near the end of a thought group; however, any syllable in the thought group can be nuclear-stressed in order to express focus by placing a pitch movement on it (Sluijter & van Heuven, 1995). The rules were presented as follows: “When no expressive stress disturbs a sequence of heavy stresses, the last heavy stress in an intonational unit [thought group] receives the nuclear heavy stress [NS]” (Newman, 1946, p. 176 as cited in Chomsky & Halle, 1968, p. 90) or as Liberman (1975) would say it, “put the strong element on the right in any given metrical constituent, if you have no good reason to do otherwise” (p. 244).

This rule (Nuclear Stress Rule) applies to both English and BP. However, the two languages exhibit subtle differences regarding this rule in wide-focus contexts. Because NS in wide-focus contexts is not the object of the present study, the next paragraphs do not mean to exhaustively discuss and explain these differences, but rather provide a general idea of how the two languages differ.

In BP wide-focus utterances, the NS falls on the last word in the utterance, be it a content or a function word (Baptista, 2001; Moraes, 1998; Truckenbrodt, Sandalo, & Abaurre, 2008), whereas in English, the position for NS is the rightmost end of the sentence, but not necessarily on the last word, as it allows for NS movement even in wide-focus contexts<sup>25</sup> (Nava & Zubizarreta, 2008; 2010). The utterances in (15) and (16) illustrate this movement in English wide-focus utterances.

(15) John read a BOOK<sub>[NS]</sub>. (Nava & Zubizarreta, 2010, p. 294)

(16) The MAIL<sub>[NS]</sub> arrived. (Nava & Zubizarreta, 2008, p. 335)

---

<sup>25</sup> This movement is ruled by the Germanic NS algorithm: “Given two sister nodes A and B in a binary metrical structure, if B is selected by A, then assign NS to B. Otherwise, assign NS to the rightmost node” (Nava & Zubizarreta, 2008, p. 335). This algorithm was first introduced in Zubizarreta (1998) and further discussed in Nava and Zubizarreta (2010).

In (15), there is a subject-verb-object wide-focus sentence, a felicitous answer to a question such as *What happened?* In this type of sentence, the NS falls on the last content word of the sentence. In (16), there is a subject-verb wide-focus sentence (an intransitive structure), which is a felicitous answer to *Why are you happy?* In this type of sentence, the NS moves to the constituent before the last.

Regarding NS in narrow-focus contexts, object of the present study, languages vary a lot: some may signal the narrow focus by means of NS (*in-situ*, that is, with no movement of the focalized constituent), while others may resort to overt syntactic or morphological cues (Elordieta, 2007). English and BP share many of these cues, as both use prosodic focus and structure focus (Frota & Moraes, 2016; Quarezemin & Neto, 2015). Both can, for example, signal focus by means of NS *in-situ* and by using syntactic cues, such as cleft sentences (Quarezemin & Neto, 2015). However, in the former, present in this study, the two languages differ.

Recall that earlier in this section (§2) it was implied that NS could be assigned to a constituent located in a place other than the final if there was a good reason for doing so (Liberman, 1975). Narrow focus as used in the present study portrays this good reason: To convey information being elicited, corrected, and contrasted. This information can be located at the end of the sentence and no NS movement is required. As a result, NS and sentence stress coincide. However, when the important information is on subject or verb position, for instance, the prominence has to move, and here lies a difference between English and BP. Zubizarreta (2016) explains that when NS and sentence stress do not coincide in location in English, NS is preserved and all sentence stresses after the NS are deleted. This phenomenon causes pitch compression after the portion with the NS and it is known as *deaccenting*<sup>26</sup> (Frota & Moraes, 2016; Gussenhoven, 1994; Ladd, 1980). Nevertheless, Gussenhoven (1983) disputes this notion of *deaccenting* in British English, by arguing that there is a pitch change (NS) at the focused constituent and another one at the end of the thought group, signaling the boundary tone. Additionally, Gussenhoven (2015) argues that the deletion of sentence stress is language dependent.

---

<sup>26</sup> *Deaccentuation* is the term more frequently used to refer to this pitch-compression phenomenon in the literature. However, in order to avoid confusion with the term “accent”, in the present study this phenomenon will be referred to as *deaccenting* (one possible term suggested by Ladd, 1980).

In BP, destressing has been found in some cases (e.g., Carpes, 2014; Fernandes, 2007a), with some pitch compression after the NS, but most research in the area reports that there is one NS at the focalized constituent and one sentence stress somewhere else in the utterance (e.g., Carpes, 2014; Frota & Moraes, 2016). Moraes (2007) states that in BP there is always a NS at the end of the thought group, irrespective of the important information being in initial or medial position, whereas a study by Fernandes-Svartman (2012) found that sentence stresses are maintained in wide-focus contexts while destressing is a possibility in narrow-focus contexts that combine NS and clefting. The next paragraphs attempt to draw more direct comparisons between NS in narrow-focus contexts in the utterance structures used in the present study: statements (NS for eliciting and correcting information) and disjunctive questions (for contrasting information).

#### 2.1.5.2.1 NS in statements

In the present study, NSs were placed in statements for eliciting information and for correcting information. In English utterances, in narrow-focus contexts, when *in-situ*, NS falls on the constituent holding the most important information in the utterance, be it in initial, medial or final position (in the latter, coinciding with final sentence stress). When in initial or medial position, that is, in a position other than the final, NS is usually followed by pitch compression (destressing) as reviewed in the previous section (Liu, 2010; Pierrehumbert, 1980). However, it is possible that more than one NS be present. The examples that follow were copied from Krifka (2007, p. 141) and illustrate this well. The (\*) signals impossible cases in English.

- (15)
- a. Which student read Ulysses?
  - b. JOHN read Ulysses. / \*John read ULYSSES.
- (16)
- a. Which novel did John read?
  - b. John read ULYSSES. / \*JOHN read Ulysses.

- (17)
- a. Which student read which novel?
  - b. JOHN read ULYSSES (and MARY read MOBY-DICK).

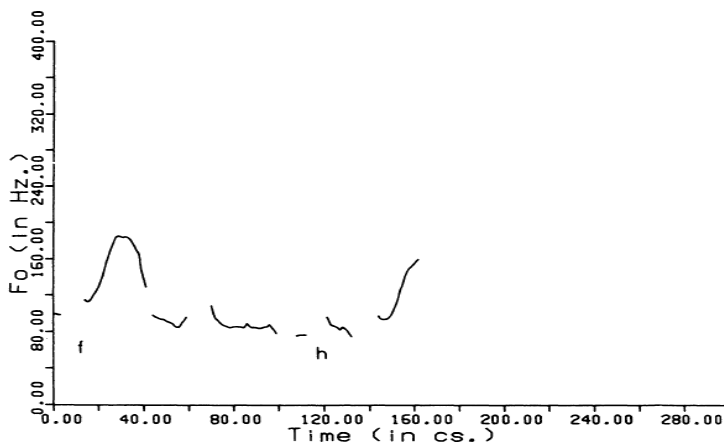
In both (15) and (16) only one piece of information is being elicited. In (15), (a) wants to know the doer of the action. Therefore, the NS lies on the subject *John* and the remainder of the utterance is destressed in (b). In (16), (a) wants to know the object of the action. Hence, the NS falls on the object *Ulysses*, which coincides with the final sentence stress, and the beginning of the utterance is destressed in (b). However, this coincidence may cause the pitch change range to be wider than if it were a pitch change of a common wide-focus final sentence stress (Büring, 2012).

In (17), (b) wants to know two different pieces of information: the doer of the action and the object of the action. Consequently, (a) uses two NSs in order to signal the two pieces of information being elicited: one on *John* and the other one on *Ulysses*. This example dialogues with Pierrehumbert (1980), who explains that two NSs in the same answer signal a felicitous answer to a question composed of two enquires. In the present study, utterances eliciting only one piece of information were used. In English, utterances correcting information follow the same trends regarding NS placement, except for one specific aspect: Pitch change range may be wider (Büring, 2012; Mioto, 2003). Nevertheless, as pointed out by Büring (2012), larger pitch range is a characteristic of narrow focus in general and the distinction between the different types of narrow focus is still inconclusive and an issue for further research.

Regarding the boundary tone, English statements usually end with a low boundary tone (L%). Nevertheless, if the discussion by Gossenhoven (1983) is considered, in British English statements that have a non-final-position NS may have a final sentence stress with a high boundary tone (H%), as illustrated in Figure 5, copied from Gussenhoven (p. 304).



(1) (It's no use calling your father)



Your father won't be able to help me.

Figure 5: *Your FATHER won't be able to HELP me.* (Gussenhoven, 1983, p. 304)

According to Gussenhoven (1983), there is a NS on *father* and some prominence on *help* that extends to the rise on *me*. He acknowledges that there is some freedom on the way this prominence (the latter) is characterized: as a NS or as a sentence stress. The present study would rather see it as the non-deleted final sentence stress, which is manifested by means of a high boundary tone (H%).

In BP statements, in narrow-focus contexts, when *in-situ*, NS falls on the constituent holding the most important information in the utterance just like in English, be it in initial, medial or final position. Interestingly, Carpes (2014) found it difficult to find in her dataset BP productions with pitch change restricted to the narrow-focused constituent when in medial position (verb). In this case, the NS was located at the rightmost end of the utterance. This may signal a difficulty for BP-L1 speakers to place NS exclusively in this utterance position.

With regard to destressing nonfocused material, when in non-final position, there is no consensus on whether NS may be followed by pitch compression with full destressing in BP. According to Frota et al (2015), when the focused information is not meant to be at the rightmost end of the utterance (i.e., in narrow-focus contexts), the material after the NS manifests pitch range compression, but no destressing. That

means that two pitch changes are produced: the NS on the focused material is manifested with a greater pitch change range while the remainder of the utterance has, at its rightmost end, another pitch change with a reduced range (which this study calls a sentence stress). However, if the focus NS is too distant from the last stressed syllable of the thought group, the pitch compression in the postnuclear position is extremely reduced that it resembles destressing, as illustrated in Figure 6, from Frota et al. However, as no measurement of pitch change is mentioned, it is not possible to tell if the pitch was compressed to less than a three-semitone range and thus be considered as non-distinctive (t'Hart, 1981).

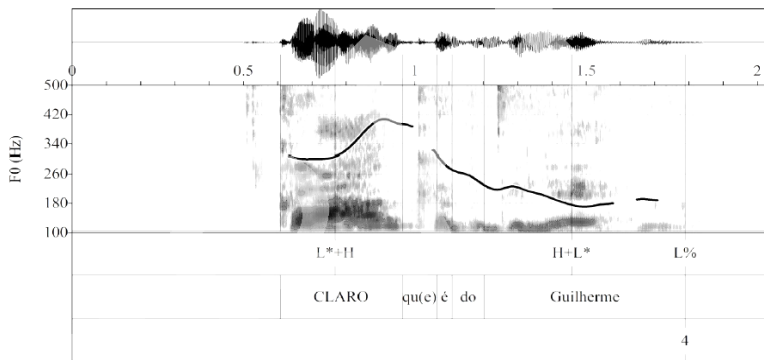


Figure 6: *CLARO que é do Guilherme* “OF COURSE it is Guilherme’s produced by a speaker from Minas Gerais (Brazil). (Frota et al., 2015, p. 26)

Moraes (1998) argues that in statements using NS to correct information, there is a pitch change at the constituent providing the correction and a pitch change on the last word of the statement. In long statements, an early emphasis has its rise prolonged until the last stressed syllable of the statement, for example, *Ele cumprimenTOU a garota de PREto* “He GREETED the girl in BLACK” (capital letters signal the pitch change and the underscore mean the maintenance of the tone). In Carpes (2014), focus in initial position showed post-focus compression, while focus in final position often had pitch changes in other parts of the utterance rather than the final only, mainly in initial position.

While in English research is nonconclusive with respect to pitch range differences between NS for information being elicited *versus* information being corrected or contrasted, in BP, research has favored

the pitch range distinction. In BP, focus in final position for information being elicited resembles that of the NS canonical position, while for correcting/contrasting information, final-positioned NS has the pitch range expanded (i.e., the rise reaches an extra-high level and the fall reaches a low level) (Frota & Moraes, 2016).

Regarding the boundary tone, BP shares with English the final low boundary tone (L%), except for the Rio Grande do Sul area, where focused statements are bounded to end with a high downstepped boundary tone (!H%) (Frota et al., 2015), which may, in a way, be compared to the British final boundary tone (H%).

#### 2.1.5.2.2 NS in disjunctive questions

One of the aims of the present study is to investigate the interpretation of the speakers' intent according to the NS assignment to signal contrasting information. These NSs were present in disjunctive questions. Syntactically, there are two kinds of disjunctive questions: *narrow-scope disjunctive questions*, which have only one interrogative clause with a disjunction, and *wide-scope disjunctive questions*, which have two interrogative clauses, conjoined by disjunction (Roelofsen & Gool, 2009). (18) and (19) illustrates them:

(18) Narrow-scope disjunctive question

Is Zenny or Anna afraid of heights?

(19) Wide-scope disjunctive question

Is Zenny afraid of heights or is it Anna?

Note that in (18), a narrow-scope disjunctive question, there is a single interrogative clause containing a disjunction, while in (19), a wide-scope disjunctive question, there are two clauses connected by disjunction. Observe that in (19), one needs to consider that some language was left unpronounced: *Is Zenny afraid of heights or is it Anna [who is afraid of heights]?* Task 2, as described in Chapter II, Section 2.3, comprised wide-scope disjunctive questions. Therefore, a review of the intonational patterns of only this kind of disjunctive question is included here.

In English, the pronunciation of disjunctive questions can be *open* or *closed*, according to the choices of the speaker. In the open

pronunciation, there is rising intonation in both clauses, while in closed pronunciation, rising intonation lies in the first clause and the second clause takes a falling intonation, as shown in Figure 7.

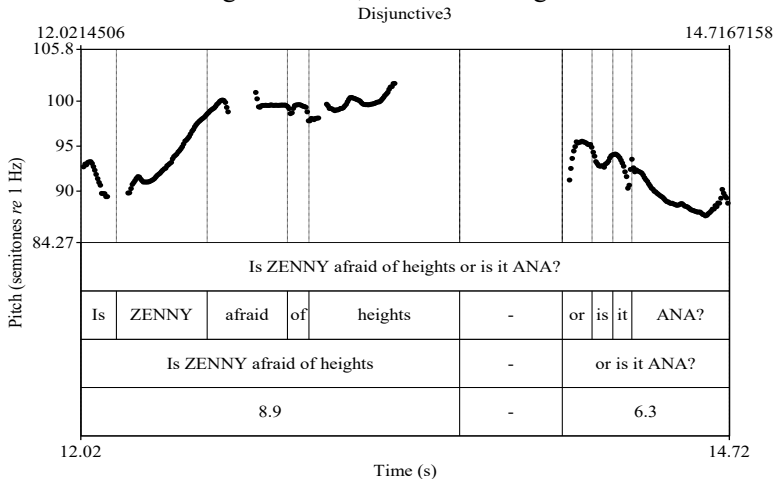


Figure 7: *Is ZENNY afraid of heights or is it ANA?* – An example of a wide-scope disjunctive question with closed pronunciation.

In Figure 7, one can see four tiers<sup>27</sup>. The first is the orthographic tier and shows the wide-scope disjunctive question with closed pronunciation. The second is the word-boundary tier and it shows the limits of each constituent in the question, while the third tier is the thought-group-and-NS tier and it displays the two thought groups for the question with the NS constituent in capital letters. Finally, the fourth is the pitch-change-range tier and shows the pitch change range in semitones. Note the rising intonation that begins in *Zenny* (the NS location), with the tone being extended to *heights*, and the falling intonation on the second clause (*or is it Anna?*), with the NS on *Anna*. In this disjunctive question, the NS falls, then, on *Zenny* and on *Anna*, signaling the contrast between the two possibilities.

The answer for the question in Figure 7 cannot be *yes* or *no*. The only possible answers would be (a) *Zenny is*, (b) *Anna is*, and (c) *Neither of them is*. Note that, in terms of pronunciation and expected

<sup>27</sup> Along the study, the Praat windows generated for the purposes of the study are described in the same manner as it is here: Tier 1: orthographic tier; Tier 2: word-boundary tier; Tier 3: thought-group-and-NS tier; Tier 4: pitch-change-range tier.

interpretation, despite of the presence of the second clause, which establishes the contrast, if the speaker misplaces the NS in the first clause, communication can get noisy and difficult. If our speaker highlighted *heights* instead, his or her listener would be expecting to hear some information other than *Anna* in the second clause, say, *big dogs*. In international communication, this can have a greater weight and impose greater frangibility, given that speakers and listeners own accents and language experience may interfere with the way they listen and interpret information. Having the NS placed at the positions which respect the contrast being made seems to be important for successful communication. In the present study, wide-open disjunctive questions with closed pronunciation were used to investigate NS placement in order to contrast information and their consequent interpretation.

BP has the same NS placements as English for wide-scope disjunctive questions (Moraes, 1998; Ribeiro, 2015). Like English, in BP there is a rising intonation (H) that begins on the NS in the first clause and a falling intonation (L) in the second clause. Nevertheless, it is not mentioned whether the high tone is maintained until the end of the first clause as it is in English when NS is in non-final positions. Due to the lack of literature describing these kinds of questions, and because of the similarity of the first clause of the disjunctive question<sup>28</sup> with the prosody and structure of polar questions, the following two paragraphs draw a brief comparison of polar questions in both languages regarding pitch compression.

In English polar questions, focus in initial and medial position has the pitch change on the focused constituent and the tone is maintained until the end of the question (a high plateau) (Liu, 2010), as illustrated in Figure 7, if one looks at the first clause of the disjunctive question only. In BP, if the question has an early focus, there is the pitch change on the focused constituent and another one in the rightmost part of the question<sup>29</sup>, as illustrated in Figure 8. Therefore, full pitch compression or destressing is not present in polar questions in BP, while it is in English, with the maintenance of the high tone.

---

<sup>28</sup> In the present study, only the production and perception of NS in the first clause of the disjunctive questions was investigated, as in Jenkins (1997).

<sup>29</sup> Note, however, that the first clause of a disjunctive question would end in an H% while the neutral polar question ends in an L%.

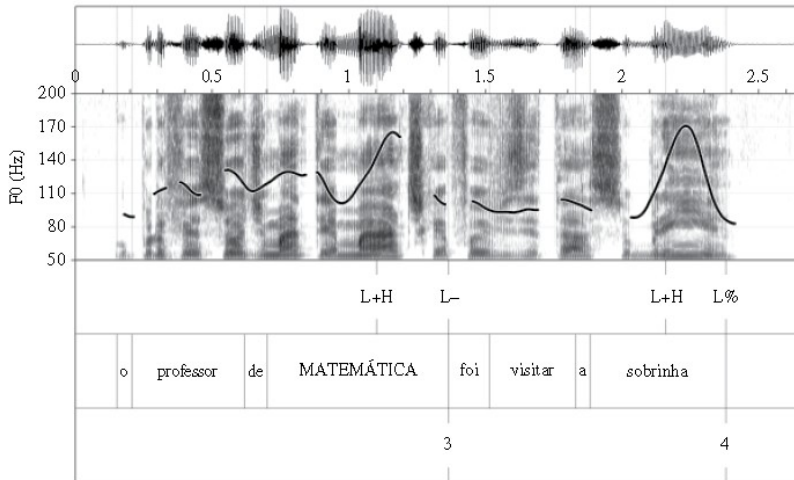


Figure 8: *O professor de MATEMÁTICA foi visitar a sobrinha?* (Has the mathematics teacher visited his niece?) (Frota & Moraes, 2016, p. 152).

Irrespective of being under narrow or broad focus, BP polar questions tend to have two sights of pitch changes, one related to the polar question intonational pattern, located at the rightmost part of the question (the sentence stress), and the other with varied pitch range, located at the leftmost part of the question (the NS). When under narrow focus, the cues for emphasis are rising movement (for initial and medial positions), gemmed pick (for medial and final positions), and longer duration (for all question positions) (Moraes, Carnaval & Coelho, 2015). The gemmed pick was also found in Mello and Silva (2015), in the use of NS for contrasting information in final position.

Table 2 summarizes the main differences and similarities regarding NS assignment in English and BP that are of interest for the present study. The next section is dedicated to talking about the unexpected assignment of NS.

Table 2  
*NS in English and BP – a comparison*

Dimension	English	BP
Wide-focus utterances	NS falls on the rightmost end of the utterance, but not necessarily on the last constituent.	NS falls on the last constituent of the utterance.
Narrow-focus utterances	Shift in NS position.	Shift in NS position.
Narrow-focus utterances	Easiness to place NS on a constituent in any sentence position.	Placing NS on a constituent in medial position is a challenge (Carpes, 2014).
Statements		
Narrow-focus statements	NS is followed by destressing. (but the exception brought by Gussenhoven, 1983)	NS is followed by pitch compression, but it is not clear if there is full destressing. Most researchers argue that there is a sentence stress at the final canonical position with a reduced range that may resemble destressing (Frota et al, 2015).
Narrow-focus boundary tones	Mostly L%; British may exhibit an H% (Gussenhoven, 1983).	Mostly L%; a high downstepped boundary tone (!H%) in Rio Grande do Sul.
Narrow-focus corrective information	No consensus about pitch range being larger than that for eliciting information (Büring, 2012).	Research has favored the claim that pitch change range is wider for information being corrected than it is for information being elicited (Frota & Moraes, 2016).
Disjunctive questions		
Pronunciation	Open or closed	Open or closed
NS assignment	On the disjoints	On the disjoints
Intonation	Rising on the first clause, falling on the second	Rising on the first clause, falling on the second
Destressing of postfocal material (based on polar questions)	A rise on the focused constituent with the maintenance of the tone until the end of the first clause.	Pitch change on the focused constituent and another one in the rightmost part of the question.

### 2.1.5.3 The unexpected allocation of nuclear stress

A frequently voiced concern is the possibility that speakers of different varieties of English will soon become unintelligible to one another. My response to such a statement is that for at least the last 200 years there have been English-speaking people in parts of the world who have not been intelligible to other English-speaking people in other parts of the world. It is a natural phenomenon when any language becomes so widespread. It is not something that is 'going to happen' but something which has happened already and will continue to occur. (Smith, 1988, p. 265)

The aspects that render speakers' speech unintelligible, as mentioned in the citation above, are manifold. The present study is mainly concerned with the effect of NS assignment, which alone may cause misinterpretation of a speaker's intent even when no unexpected pronunciation of segments is existent (Galaczi et al., 2017).

The principles guiding NS placement in English and BP coincide in some moments while they are alike in other moments, as reviewed in Section 2.1.5.2. The main difference lies on the fact that in BP, when in wide-focus contexts, NS is placed in the last word of a thought group, irrespective of being a content word or a function word. To illustrate it, consider an often-cited example provided by Baptista (2001), in (20).

(20)

- a. Dê o livro pra MIM. (wide focus in BP)
- b. Give the book to ME. (narrow focus in English)

In BP, the NS on *mim* is interpreted as non-contrastive new information, that is, it is considered to be in wide focus and the whole sentence is considered new information. Therefore, NS falls on the canonical position for PB: The last constituent in the sentence. It is a request that can perfectly be uttered in a room occupied by both speaker and listener only or by the two accompanied with more people. In English, however, if the NS location is directly transferred from the speaker's L1 to English, it signals contrasting information (narrow-focus



contexts). If nobody else were in the room, the listener would probably wonder the reasons why such a request was being made, and if there were somebody else, the listener would go suspicious that something else would be going on, and thus misinterpret the speaker's intent.

Jenkins (2000) describes an interaction among four students from different nationalities (Brazilian, Swiss-French, Colombian, and Hungarian), while making posters for the classroom wall. The Hungarian student asks the other three the question in (21).

(21) Have you got a blue VUN?

The scholar reports that the three other interactants echoed the words *blue vun* and *vun* many times and got the intended meaning only after the Hungarian student, holding up a blue pen, explained: *Blue vun like THIS*. The intended meaning for *vun* was *one*. The author highlights that, although interactants were acquainted with each other accents in English and had enough contextual cues in order to provide the listeners with clues to meaning (they were making posters and surrounded with paper and colored pens), the misallocation of NS was a great villain to the lack of success in communication in that interaction. Jenkins argues that the mispronunciation of *one* would have caused no problems if the NS had been placed in *blue* (i.e., as opposed to *red*). Once it was placed in *vun*, no contrast was possible, and it signaled that *vun* carried the most important information in the thought group, misleading the interpretation of meaning.

The inability to segment speech into meaningful thought groups is one of the most common factors contributing to problems with NS placement (Jenkins, 2000). Celce-Murcia et al. (1996) explain that the number of NS in a given utterance depends on the speaker. The more pauses the speaker produces, the more thought groups the speaker creates, and thus, the more NS allocations arise. They highlight that too many nuclear stresses (due to many pauses) make the overall message difficult to process and understand, which may lead or not to unsuccessful communication. Excessive pauses were present in the dataset of the current study, as reported in Chapter 4.

### 2.1.6 Summary

This section defined some important concepts for the present study. Language is seen as a complex adaptive system, and thus, speech communication is also complex and dynamic. English is seen as an International Language due to its wide use by speakers around the globe. Due to the different L1 background of IL-English speakers (the notion of strange attractors) and to the chaotic nature of human language itself (the unexpected and unpredictable changes), associated with some other factors (e.g., the advances in technology), variation and accent is everywhere. This ubiquity of accent may enhance the chances of miscommunications around the globe. Making use of accommodation strategies is one of the ways to avoid this problem, as speakers converge towards the speech of their interlocutors in international contexts mainly in order to improve interpretability (Jenkins, 2000), which encompasses three dimensions: intelligibility, comprehensibility, and interpretability (Smith & Nelson, 1985). However, because of repertoire limitations, in face-to-face interactions, a motivated-to-converge IL-English speaker may lack the ability to do so due to lack of proficiency (Jenkins, 2000). Therefore, it is very important that an individual is equipped with the phonological features that are necessary to accommodate both as a speaker and as a listener. The present study aims at contributing to informing pedagogy on the phonological repertoire that is a must to enhance BP changes to succeed in communication, more specifically with regard to the use of NS, the assignment of more prominence to a constituent that holds the most important piece of information in a sentence. NS assignment has some coincidences in English and PB, but the two languages differ mainly with regard the canonical position in wide-focus contexts and post-focal destressing in narrow-focus contexts mainly in polar questions. The unexpected placement of NS may cause the speaker's intent to be misinterpreted, and the production of more than one NS in a sentence breaks it into many thought groups, which may make the overall message difficult to attain. The next section reviews some studies regarding the perception of NS and the speakers' intent interpretation.

## 2.2 THE PERCEPTION OF NS AND THE INTERPRETATION OF SPEAKERS' INTENT

### 2.2.1 Cues for NS perception

The perception of NS depends on acoustic cues such as pitch, loudness and duration (Calhoun, 2007) and it is argued that the narrower the focus the more identifiable a NS (Ortiz-Lira, 1998).

Portions of speech in which NS falls tend to have extra length and loudness. The vowels in the NS syllable have fuller vowel quality than those in the other portions of a given thought group, the duration of both stressed and unstressed syllables is lengthened, and its intensity is also affected (Sluijter & van Heuven, 1996b).

Sluijter and van Heuven (to appear, as cited in in Sluijter & van Heuven, 1996b), investigated stressed and unstressed vowels in syllables with and without a NS using American English minimal stress pairs and their reiterant speech-copies. The *F0* contours and formant values were determined and the spectra were examined. Results showed that (1) NS-lending *F0* movements occurred only on focused targets and non-focused targets were produced with minor *F0* changes; (2) regarding quality, stressed vowels were fuller than unstressed vowels, and stressed vowels holding a NS were fuller than the regular stressed vowels; (3) glottal pulses were more sinusoidal in unstressed syllables, that is, high-frequency emphasis was weaker, indicating smoother and slower vocal fold closing movement; (4) high-frequency was higher in NS vowels than in regular stressed vowels; and (5) glottal leakage<sup>30</sup> was larger for stressed than for unstressed vowels, and there were insignificant differences regarding word stress and NS.

Sluijter and van Heuven (1996b) compared the acoustic correlates of stress with the intervening effect of sentence stress in American English. The correlates investigated were (1) *F0* (fundamental frequency), (2) duration, (3) overall intensity, (4) source parameters (related to spectral balance: open quotient, amplitude of voicing, closure rate/skewness of the glottal pulse, and glottal leakage), and (5) filter

---

<sup>30</sup> Glottal leakage, when constant, causes the production of audible friction noise. It usually happens due to the incomplete closure of the glottis. For more information, access <http://www.ims.uni-stuttgart.de/institut/arbeitsgruppen/phonetik/EGG/page10.htm>.

parameters ( $F1$  and  $F2$  as acoustic correlates of vowel quality). They analyzed a recorded speech corpus of six L1-American-English speakers containing four noun-verb minimal stress pairs and three different reiterant speech-copies, with and without narrow-focus NS in fixed carriers. Their results imply a hierarchy of cues to word stress: (1<sup>st</sup>) duration, (2<sup>nd</sup>) glottal parameters (i.e. high-frequency emphasis and glottal leakage), (3<sup>rd</sup>) vowel quality, (4<sup>th</sup>) fundamental frequency, and (5<sup>th</sup>) overall intensity. The two latter showed to bring in little or no cues to word stress position. As to NS, fundamental frequency, overall intensity, open quotient, and amplitude of voicing were found to be reliable correlates.

Shue, Iseli, Veilleux, and Alwan (2007) investigated the acoustic correlates of NS and word stress in American English, and the interaction of these correlates with other factors that affect prosody. The variables were presence/absence and type of NS ( $L^*$  or  $H^*$ ), boundary-related tone sequence (LL% or HH%), and gender of the speaker. The measures were duration, fundamental frequency ( $F0$ ),  $H^*_1 - H^*_2$  (related to open quotient), and  $H^*_1 - A^*_3$  (related to spectral tilt). Five males and five females produced 10 times each of the four versions of a sentence (*Dagada gave Bobby doodads*). The versions differed in the placement of NS (initial and final) and whether it was a statement or a question ( $L$  and  $H$  boundary tones). They found that (1) nuclear stresses were clearly marked by differences in  $F0$  contours, (2) syllable duration was longer for the NS cases than for non-NS cases, (3) on stressed syllables, there was a tenser voice (lower open quotient). In all, results mainly suggest that acoustic cues of word stress can be affected by the presence of a NS, boundary tone, and, not in all cases, gender of the speaker.

The studies reviewed in the paragraphs above indicate that stressed syllables in non-focused material are not as prominent as stressed syllables in focused material, be it in wide-focus or narrow-focus contexts. Previous literature has found evidence that NS in narrow-focus contexts are more prominent than those in wide-focus contexts (Calhoun, 2007; Klok et al., 2011). However, there is no consensus whether NS prominence in narrow-focus contexts varies according to the discursive function (e.g., for correcting and eliciting information) (Büring, 2012).

A study within this vein was conducted by Klok et al. (2011), who compared the use of focus in English (35 L1-English speakers), Québec French (17 L1-Québec French speakers), and European French

(16 L1-European French speakers) in four types of contexts: (1) parallelism<sup>31</sup>, (2) contrast, (3) correction, and (4) control (wide-focus context), by analyzing acoustic duration, intensity, and pitch cues. Their results showed that English signals the types of focus investigated by means of duration, intensity and pitch change, while French, both from Québec and Europe, uses only duration and pitch change, and fails to signal parallelism. They also found that there are “differences in how a prominence shift is implemented depending on focus: in English, in contrastive focus and parallelism, the prominence is shifted by reducing duration and intensity on the given word, while in corrective focus the contrastive word is boosted. With respect to pitch, however, the contrastive word is boosted in both corrective and contrastive focus, while the pitch of the given word is lowered in the case of parallelism.” (p. 1). Their study supports the assumption that NS to contrast and correct information have more prominence than other types of NS uses (in their case, when compared to parallelism and wide focus).

According to Calhoun (2007), not only acoustic cues are important for the perception of prominence, but also is the position of prominence in prosodic structure. If the focused material is placed in the rightmost part of the utterance where NS is already expected, its prominence has to increase gradually so it can be perceived. In her study, Calhoun (2007) found that for words that are expectedly stressed, such as nouns or verbs, prominence has to increase even so they are focused. Pronouns, on the other hand, are usually unstressed and if they have some prominence they are likely to be focused. The author highlights that both acoustic and structural prominence are manipulated to convey focus and, although both pitch and duration increase to convey focus, “duration is a better cue than pitch to prosodic structure.” (p. 624) The author argues that the amount of prominence of focused material is closely related to the availability of contrasting alternatives in the utterance: the more prominent the focused material, the more available its alternatives.

Despite the importance of acoustic cues and sentence structure, Bernstein et al. (1989, as cited in Keating et al., 2003, p. 2071) found that NS can be perceivable from visual-only speech. Keating et al. (2003) set out an investigation to scrutinize the optical phonetic

---

<sup>31</sup> In their study, an example of parallelism was “Yeah, yesterday, he fixed up a *blue bike* and a **red bike**”, a felicitous response to the stimuli “I heard that Jordan is into cycling” (Klok et al., 2011, p. 1).

characteristics that allow visual perceivers to recognize stress. In articulatory terms, a NS is often associated with larger, longer, and faster articulations. By drawing comparisons between visual cues for word stress and NS, the scholars found that (1) NS cues abound more than those of word stress (e.g., lexical stress was not associated with eyebrow movement) and (2) perception results were comparable to those of production, as participants perceived NS more accurately than word stress. They concluded that larger and faster mouth opening movements, more open mouth positions, and head movements are visual cues for both word stress and NS, while eyebrow movement is more related to the latter.

In the present study, NS was mapped mainly by means of pitch tracking on *Praat* (Boersma & Weenink, 2014) and perceived loudness by the researcher and listeners. Additionally, only an impressionist account of duration was made as constituents varied in length to serve the purposes of the present study. Regarding visual cues, speakers' movements of the head and eyebrows were mapped in order to help understand intriguing and inexplicable expected perceptions by the listeners. The next section reviews studies about NS worldwide.

## **2.2.2 Production and perception of NS and speakers' intent interpretation worldwide**

Studies on the intelligibility of IL-English speakers abound. However, most studies have investigated the dimensions of intelligibility and comprehensibility only, leaving the interpretability of speakers' intent aside. Smith and Nelson (1985) review state-of-art information on intelligibility within the period of 1950-1985 gathered from 163 pieces of research. The authors mainly claim that future studies should make distinctions between intelligibility, comprehensibility, and interpretability. Moreover, the authors summarize the most important agreements on the field, which are depicted as follows:

- 1) The existence of IL-English speakers with unintelligible English is a natural phenomenon. It is said that one has to be intelligible only to those with whom one will interact.
- 2) L1-English speakers (a) are not the only judges of what is intelligible in English and (b) are not more intelligible than IL-English speakers.

- 3) Intelligibility is interaction-dependent, and influenced (a) by experience with either the interlocutor's accent or the interlocutor's language (belonging to a given speaking community), and (b) by the listener's expectations on (positive attitude and willingness to understand) the interlocutor's speech.

Attention should be drawn to the first agreement mentioned. Perhaps at that time (1985) it would seem right. However, thinking this agreement over may raise the question of how one can possibly tell who one's interlocutors are supposed to be in the course of life. Nowadays, chances of interaction with people from different parts of the world are high, made possible mainly through the use of social medias (e.g., Facebook®) in the Internet. Thus, with regard to pronunciation, a speaker should work on the features that are a must to guarantee his/her intelligibility when interacting with both L1 and IL-English speakers. One of the objectives of the present study is discussing the role of NS for Brazilians' communication with people around the globe in English.

Regarding the second consideration, more updated studies have also shown that L1-English speakers are not more intelligible than are IL-English speakers (e.g., Atechi, 2004). A possible explanation for this is the amount of resources available for the listener, that is, the more proficient a listener is, the more complete his/her phonological system is in English as an IL, and thus, the more language s/he is able to recognize and thus understand. The present study takes into account proficiency as a factor for both speaker and listener.

Smith and Nelson (1985) finish the article by pinpointing some aspects worthy researching. They are:

- 1) Correlation of (a) proficiency and topic difficulty of both listener and speaker to (b) the comprehensibility, intelligibility, and interpretability of the speaker's speech.
- 2) Communicative setting influences.
- 3) Further research on familiarity, speaker and listener efforts to communicative success and attitude towards different varieties of spoken English.
- 4) Research involving (a) IL-English speakers as informants and (b) other IL-English speakers from different nationalities as listeners.

- 5) Research involving (a) L1-English speakers as informants, (b) IL-English speakers and (c) L1-English speakers from different countries as listeners.

The present study responds mainly to aspect four, provided that its design includes (a) BP-L1 speakers of English as informants and (b) IL-English speakers from different nationalities as listeners. Additionally, an L1-English speaker participated as a listener, since the chance of interactions between Brazilians and L1-English speakers are also high<sup>32</sup>.

Although follow-up studies have revolved around some of the issues pinpointed by Smith and Nelson (1985), much research is still needed to reach an understanding of the intricate complexity of communication among different speakers of English. The paragraphs below review some research conducted, more specifically in the field of prosody.

Languages differ with regard to the use of NS and some research has been dedicated to understanding if NS rules are transferred from the L1 to the IL-English speakers' speech. Nava and Zubizarreta (2010) compared the production of L1-English and L1-Spanish speakers of English and hypothesize that Spanish speakers would transfer the NS rule from their native language. Participants in the study were 54 speakers living in Los Angeles divided into two groups: 30 L1-English speakers (control group), and 24 L1-Spanish speakers of English (experimental group). In the control group, 12 had an intermediate level of proficiency while the other half had a high level of proficiency in English (tested via a Cloze test). Their hypotheses were confirmed in two cases: (1) the case of NS placement in SV (subject-verb) unaccusative structure, in which the L1-Spanish English speakers placed the NS mostly on the verb while the L1-English speakers placed NS on the subject, and (2) the case of English compound OV (Object-Verb) structure, in which NS was placed mostly on the object by the L1-English speakers and on the verb by L1-Spanish English speakers. However, the other structures (SAdvV and SVAdv) investigated yielded no differences in terms of NS placement: both groups placed NS on the last word of the sentence, that is, on the verb in the case of the SAdvV

---

<sup>32</sup> In fact, the present study takes the position that L1-English speakers besides using English for communicating with other L1-English speakers also use English to communicate with the international community. Consequently, L1-English speakers cannot be excluded from the large community that uses English to communicate internationally.



structure and on the adverb in the case of the SVAdv structure. Another hypothesis was that L1-Spanish speakers would acquire destressing before they acquire the NS rules for English and it was confirmed. The authors concluded that “L2 learners start out computing NS on the basis of the general NSR (...). In addition, (...) learners acquire A-deacc [destressing] before they acquire the Germanic NSR” (Nava & Zubizarreta, 2010, p. 316).

Harris (2014) compared the prosodic marking of given information in English and Italian, produced by five L1-English and five L1-Italian speakers in their own languages and by an advanced L1-Italian speaker of English in English. Their productions were triggered by means of a card game, which repeated referents within noun phrases. Results showed that L1-English speakers signaled old information via stress distribution and duration, while L1-Italians’ stress distribution and use of duration mismatched signaling given information as they stressed and had similar durations for both new and old information. Equally, this lack of match between these prosodic cues and old information was found in the productions of the L1-Italian speaker of English. Based on these results, the scholar suggests that prosodic transfer is persistent at advanced levels.

Na, Yuan, and Bin (2012) investigated how the production of narrow focus in English by eight L1-Zhenjiang-dialect speakers of English compare to that of six L1-English speakers. Participants recorded carrier sentences containing two, three and four-syllable words (“I said \_\_\_\_\_ ten times”). Results showed differences in terms of reduction of unstressed material and peak alignment, mainly in the case of sentences containing words with stress on the third syllable. L1-Zhenjiang-dialect speakers also placed NS on words other than the ones in the focus domain, and, some placed NS on a word other than the word in narrow focus. This dialect shares with PB the lack of fully destressing postfocal material and the multiple unexpected NS found in their productions might have been a result of language transfer.

NS placement has shown to interfere not only with the comprehension of the message but also with the way speakers’ are judged by listeners. Hahn (2004) investigated the reactions of 90 L1-English speakers to speech of one L1-Korean speaker of English, which included samples with NS correctly placed (stimuli type 1), incorrectly placed (stimuli type 2), or missing entirely (stimuli type 3), that is, when the speaker preserved or violated the *old-new stress connection*.

Listeners were randomly assigned to three groups and each listened to only one type of stimuli (1, 2, or 3). Listener's difficulty processing the discourse in each text version was assessed by using the dual-task paradigm borrowed from psycholinguistic studies of spoken language processing, comprehension was assessed through a write-down-as-much-as-you-understood task and a short-answer comprehension quiz, and evaluative reactions to the speaker was verified by a pool of Likert-type items usually used for collecting summative feedback about classroom instruction. Results showed that participants who listened to stimuli *Type 1* tended to perform more quickly the concurring task if compared to those who listened to stimuli *Types 2* and *3*, but differences were not statistically significant. Regarding the comprehension of discourse, results indicated that participants that listened to stimuli *Type 1* recalled more main ideas and more details than the other two groups; differences were not significant for details recalled, though. Finally, participants who listened to stimuli *Type 1* evaluated the speaker more positively than those in the other two groups and differences were significant. The author concluded that participants generally responded more positively to the discourse of the IL-English speaker when the speaker preserved rather than violated the old-new stress connection and that although differences were not always significant, results showed to follow a similar pattern, one that supports the general proposition that expected NS placement by IL-English speakers facilitates communication. Hahn (2004) explains that oral texts that lack NS allocation apparently convey the impression that the speaker is speaking rapidly and thus keeping track of what is being uttered is more challenging/problematic to listeners.

Zoghbtor (2010) investigated how effective a pronunciation syllabus based on the LFC (by Jenkins, 2000) was in improving the intelligibility and comprehensibility of 50 Arab learners. Participants were divided into two groups of 25 learners: experimental (receiving the LFC pronunciation syllabus) and control (receiving traditional pronunciation syllabus). The experimental group gain scores were higher than were those of the control group, but differences did not reach statistical significance. As to the placement of NS, the scholar found that it is closely related to the comprehension of the message rather than to the recognition of words and that it is important not only to facilitate comprehension and intelligibility, but also to trigger a positive judgment over a speaker's speech. Additionally, speakers

allocating NS as expected were seen as more connected to the message being conveyed, and thus as more interesting to be listened to, as reported by the L1 and IL-English listeners that participated in the study.

Another study guided by the Lingua Franca Core principles (Jenkins, 2000) was conducted by Andrade (2005), who investigated the performance of 40 IL-Japanese speakers of English on a perceptual test of English syllabification, word stress, NS, and thought groups. Five tests were administered: (1) a syllabification test, whereby learners had to parse the words by analyzing both written and recorded stimuli; (2) a word stress test, whereby learners had to circle the stressed syllable after listening to each word; (3) a sentence stress test, whereby learners had to underline the words with prominence after listening to each mini-dialogues; (4) a NS test, whereby learners had to distinguish differences in meaning when the focus of a sentence changed, by choosing, after listening to stimuli, one out of five options in their sheet; and (5) a thought group test, whereby learners had to distinguish differences in meaning based on pauses, linkage, juncture, and other thought-group-related variables. The results indicated that participants were moderately good at distinguishing meaning between paired sentences with shifts in NS (Test 4), which corroborates Jenkins' proposition that learners have the receptive skill for NS, while they showed difficulties in parsing, word stress, sentence stress, and thought group identification.

Some studies have been conducted to investigate the interpretation of the speakers' intent and message when relying on the perception of NS. A study carried out by Tiffen (1974) investigated the intelligibility of educated 24 L1-Nigerian-English speakers to L1-British-English listeners in both segmental and suprasegmental levels. For the latter, word stress, NS, and intonational patterns were analyzed. For the NS investigation, speakers had to stress different portions of a sentence according to the stimuli given by the researcher via a context question as illustrated in (22).

(22)

Interviewer: Did BILL motor to London?

Speaker: No, JOHN motored to London. (i.e., not BILL)

Interviewer: Did John CYCLE to London?

Speaker: No, John MOTORED to London. (i.e., not CYCLED)

Interviewer: Did John motor to MANCHESTER?

Speaker: No, John motored to LONDON. (i.e., not MANCHESTER)

Production results showed that all the L1-Nigerian-English speakers had difficulty with this feature of English pronunciation ( $M = 40.4\%$ ), placing NS in the rightmost portion of the utterance, irrespective of the correction being made. Some of the productions were randomly selected to compose the listening task taken by the L1-British-English speakers. Due to the unexpected NS assignments in the production level, listeners were guided by the perception of the unexpectedly assigned NS and thus failed to identify which correction was being made and misinterpreted the speakers' intent.

In line with Tiffen (1974), Atechi (2004) investigated mutual intelligibility of L1-Cameroon-English speakers with L1-British-English and L1-American-English speakers in both segmental and suprasegmental levels. For the latter, both word stress and NS were examined. The NS procedures and instruments resemble the ones used in Tiffen's study. It differs only in that stimuli were produced by the three groups of speakers, and heard by listeners from the same three nationalities, viz. Cameroonian, British, and American. Production results showed that Cameroonian speakers failed to place the NS in order to highlight the correction intended, and, accordingly, listeners (both British and American) failed to interpret the speakers' intent.

Lanham (1984) investigated the consequences of unexpected allocation of stress at both word and sentence levels (word stress and NS) by a South-African-Black-English speaker (whose L1 was described as the Bantu languages of southern Africa). Listeners were 13 L1-white-English speakers and two L1-South-African-black-English speakers. Participants listened to the recorded passage and answered a comprehension quiz assessing the comprehension of the passage read by the South-African-Black-English speaker. Regarding NS, results showed that the misallocation of NS (in order to establish focus on new information) was cause of interruption of syntactic content and thus

posed difficulties for the listeners in making sense of the message being conveyed.

A piece of research that is of particular interest for the present study was conducted by Jenkins (1997). In her text, Jenkins argues that discourse intonation is problematic for two main reasons. The first relates to the dichotomy of old/new. The author argues that it is clear but at the same time obscure, since what is new or old depends on many factors that are not present in the context of discourse. However, Jenkins does not explore these factors to support her ideas<sup>33</sup>. The second reason lies on the teachability of pitch. She argues that pitch patterns become unnatural or even wrong when produced by teachers. According to the scholar, a specific trait of intonation, however, is more likely to be both teachable and learnable: the accentual function of intonation (NS). She claims that NS placement “operates at a more conscious level than the other aspects of the intonation system” (p. 18). NS, in wide or narrow-focus contexts, is the most important feature to signal the speaker’s intended meaning, mainly because it relies on stress to highlight contrasts, since word order in English is relatively inflexible (Jenkins, 1997). (23), (24), (25), and (26) illustrate frequent unexpected NS uses mapped by Jenkins.

(23) I smoke more than you DO. (Jenkins, 1997, p. 18)

(24) The black COLour, not the blue COLour. (Jenkins, 1997, p. 19)

(25) Do you have a blue VUN? (Jenkins, 1997, p. 19)

(26) Do you have to use SOMETHing for the wasting time? (Jenkins, 1997, p. 19)

In these instances, speakers produced “the same type of mistake, either by misplacing *unmarked* [emphasis added] [wide-focus] nuclear stress, or more commonly, using unmarked nuclear stress where *contrastive* [emphasis added] [narrow-focus nuclear] stress was required for the meaning” (Jenkins, 1997, p. 19). Despite not being able to use NS to establish contrast, these speakers relied on NS to interpret

---

<sup>33</sup> The notion of *new* in the present study was established in Section 2.1.5.1.

messages, which may indicate that they have the NS receptive skill (which was later corroborated by Andrade, 2005).

In order to disentangle the difference between reception and production of NS, Jenkins (1997) conducted a study with eight participants divided into two groups: experimental, with four IL-English speakers, and control, with four L1-English teachers with little phonological awareness.

The dataset consisted of four sets of five questions that required the use of NS in narrow-focus contexts. The first clause of a disjunctive question in each set was the same and had five different second clauses, as illustrated in (27) (Jenkins, 1997, pp. 19-20).

(27)

Did you buy a tennis racket at the sports center this morning, or...

- a. was it a squash racket?
- b. did you buy it yesterday?
- c. did you only borrow one?
- d. was it your girlfriend who bought it?
- e. at the tennis club?

A set was produced by only one speaker in each group (control and experimental). Speakers were asked to circle the words they meant to highlight. So, in all, 40 questions were recorded and, after having the second clauses removed, they were submitted to the two groups of participants as a listening task. In this listening task, participants in the two groups (control and experimental) were asked to listen to the first clause and predict the second clause in each question. Data analysis showed that the control group (L1-English speakers) placed NS on the expected portions of the first clauses while the experimental group (IL-English speakers) did not. As for the listening task, the control group predicted correctly all the second clauses and the experimental group predicted correctly two thirds of the second clauses. Results of Jenkins's study supported her hypothesis that learners acquire NS at the receptive level relatively fast, while at the productive level it is not the case. These results support the importance of overtly teaching NS in language classrooms.

From 1950 up to now, a great body of research on IL speech intelligibility has been conducted. Nevertheless, much more is needed in

order to help us understand the complexity of English as an IL communication. The next portion of this section reviews some studies conducted in Brazil.

### **2.2.3 Production and perception of NS and speakers' intent interpretation in Brazil**

In Brazil, intelligibility has been investigated to inform pedagogy mostly at the segmental level (e.g., Becker, 2011; 2013; Cruz, 2003, 2004, 2008; Schadech, 2013; Schadech & Silveira, 2013). A few other studies have investigated the effect of non-target production of suprasegments on the intelligibility of BP-L1 speakers of English (Gomes, Brawerman-Albini, & Engelbert, 2014; Marta, 2011; Passarella-Reis, Gonçalves & Silveira, 2016), but, to the best of my knowledge, research on the use of NS associated to interpretability of BP-IL speakers' speech in English is nonexistent. The next paragraphs review four studies that have tackled the production and awareness of NS in English utterances by Brazilians.

De Castro Gomes (2012) reports on the results of manifold analyses conducted by a study group on the Brazilian way to speak English. Regarding NS, a comparison was drawn between the production of NS by three L1-English speakers (two Americans and one British) to that of three L1-BP advanced speakers of English. The analysis of data for the production of *Where did all this money come from?* showed that the L1-English participants consistently placed NS on the constituent *money* (utterance non-final located) while the Brazilian participants assigned NS to the constituent *from* (utterance final located), following the NS canonical position in PB. The results reported indicate that these Brazilians tended to transfer the NS L1 use to English even at advanced levels of proficiency and corroborates the findings of studies conducted with other languages (e.g., Harris, 2014).

Manosso (2013) investigated the assignment of NS by 84 L1-BP speakers of English (who differed in level of proficiency, length of formal English instruction, and length of residence in an English-speaking country) and four L1-English speakers<sup>34</sup>. Participants recorded the reading of a text in English, out of which one question was considered in the data analysis: *Where did all this money come from?*

---

<sup>34</sup> Some of Manosso's (2013) data were from the study reported by de Castro Gomes (2012).

Although statistical tests were not run, the analysis showed that (1) NS was placed in the L1-English location more often by the more proficient L1-BP speakers and by participants who had lived in an English-speaking country longer than six months, and that (2) length of formal English instruction did not play a role. The author also found some movement in the leftmost part of the question (*Where*), but whether or not it would be a NS was left for future research. This initial pitch change is a trait of BP language in polar questions (Moraes, 2015).

Passarella-Reis and Silveira<sup>35</sup> (2016) investigated NS placement by L1-BP speakers of English (speakers) while interacting with other L1-BP speakers of English (listeners). Speakers engaged in a task in which they had to read aloud an utterance in response to a context question or a context sentence and assign NS according to the discursive context. The analysis of data showed that the speakers had difficulties (1) to place NS as expected, irrespective of utterance position (initial, medial or final) and (2) to signal information being corrected (corrective focus) and elicited (presentational focus) by means of NS. The unexpected NS assignments were classified into three categories: (1) the assignment of only one NS on a position other than on the expected portion of the utterance; (2) the assignment of two or more NSs, being the one placed at the expected portion of the utterance the NS holding a greater pitch change range; and (3) the assignment of two or more NSs, with no distinct differences of pitch change range between them. The authors highlight that the unexpected NS use by these participants may compromise the way these speakers' intent might be interpreted during their interactions in English.

Kivistö-de-Souza (2017) investigated the perception of NS by 69 L1-BP learners of English and 16 L1-English speakers. Data was collected by means of a perception task that presented low-pass filtered utterances varying in NS assignments, most of which had wide information focus (75%). Utterances also had three structure types: (1) ending in old information, (2) ending in relative clauses, and (3) ending in function words. Participants had to decide if the intonation was appropriate or not for a given utterance according to a given context. The scholar found that L1-BP learners of English were less aware of NS assignment than were L1-English speakers. Another finding was that NS perception was affected by type of utterance: the perception of NS was

---

<sup>35</sup> Passarella-Reis and Silveira (2016) report on the production results of the study conducted to pilot the present study.



higher in destressed sentences than in unaccusative sentences. Additionally, it was found that L1-BP speakers of English showed higher awareness about NS assignment in utterances ending in function words than in old information. The latter finding indicates that L1-BP speakers of English that participated in the study had not developed awareness about NS assignment in utterances ending in old information. She concluded that the L1-BP learners' awareness about NS placement was limited (to function words only) and she suggested, by drawing on literature (e.g., Zubizarreta & Nava, 2011) and on the results of her study, that L1-BP learners of English would benefit from explicit prosodic instruction even at advanced levels of proficiency.

The four studies reviewed above do not approach the way Brazilians' intent is understood when they use NS to signal important information. The present study aims at contributing to filling this gap.

#### **2.2.4 Summary**

This Section reviewed studies that investigated the cues for the perception of NS, NS transfer from L1 to English, the production and perception of NS and the consequent interpretation of speakers' intent and message being conveyed.

Regarding the cues, studies agree that the NS constituent presents great pitch changes, but no consensus has been reached about if the range of these changes varies according to the function of NS. Besides the acoustic cues (e.g., duration, vowel quality), visual cues may also contribute to the perception of NS, mainly eyebrow and head movements.

Studies on the production, perception and interpretability of speakers' intent based on NS do not abound in the literature worldwide and in Brazil. The studies reviewed here have suggested that IL-English speakers tend to transfer the NS rules from their L1 to English (e.g., Harris, 2014; de Castro Gomes, 2012), assign NS to unexpected portions of utterances, which compromise the way they are interpreted (e.g., Tiffen, 1974), do not destress given material and thus produce a greater number of NS, which makes it difficult for listeners to understand the main idea of what they are saying and makes their speech less pleasant to be heard (e.g., Zohbor, 2010). In Brazil, the main findings about NS assigned in English by Brazilians are: (1) they are not fully aware of NS use in English, mainly with regard to destressing old information

(Kivistö-de-Souza, 2017); (2) they transfer their NS use from BP to English (de Castro Gomes, 2012; Manosso, 2013); and (3) they have difficulties to assign NS in accordance with a given discursive context, that is, NS is placed in more than one constituent irrespective of utterance position and meaning (Passarella-Reis & Silveira, 2016).

### 2.3 CONCLUSION OF THE CHAPTER

This chapter was organized into two previous sections and the present one. Section 2.1 was dedicated to defining the most important concepts for the present study: (1) language, which is seen as a dynamic adaptive system (Larsen-Freeman, 1997) (Section 2.1.1); (2) the status of English as an International Language (IL) (McKay, 2010) and who the speakers of English as an IL are (Section 2.1.2); (3) communication, interaction – seen as dynamic and adaptive processes as language itself is – and Communication Accommodation Theory, which makes available strategies for face-to-face interactions (Section 2.1.3); (4) speech communication (Catford, 1950) and speech intelligibility (Smith & Nelson, 1985) (Section 2.1.4); (5) nuclear stress to signal the most important part of the utterance (focus) and the terminology involved in prosody, such as sentence stress, thought group and intonation (2.1.5). Additionally, Section 2.1.5 discussed the function of NS (Section 2.1.5.1), compared NS placement in English and in BP (Section 2.1.5.2) and discussed the unexpected allocation of NS (Section 2.1.5.3).

Section 2.2 reviewed studies on the cues for the perception of NS (Section 2.2.1), production of NS by the international community and by Brazilians as well as the interpretation of speakers' intent based on NS placement (Sections 2.2.2 and 2.2.3).

Most of the studies reviewed in Sections 2.2.2 and 2.2.3 were conducted to investigate only the production of NS (de Castro Gomes, 2012; Harris, 2014; Manosso, 2013; Nava and Zubizarreta, 2010; Na et al., 2012; and Passarella-Reis & Silveira, 2016), only the perception of NS (Kivistö-de-Souza, 2017), the production and perception of NS and the consequent interpretation of speakers' intent (Atechi, 2004; Jenkins, 1997; and Tiffen, 1974). On the one hand, the present study's purpose equals Atechi's (2004), Jenkins's (1997) and Tiffen's (1974) as it means to investigate both the production and the perception of NS and the consequent interpretation of speakers' intent. On the other hand, it

differs from these studies regarding the participants involved and the research design.

Participants in the studies reviewed usually are IL-English speakers of a specific nationality and L1-English speakers (de Castro Gomes, 2012; Hahn, 2004; Harris, 2014; Kivistö-de-Souza, 2017; Manosso, 2013; Na et al., 2012; Nava & Zubizarreta, 2010). Participants in the present study are L1-BP speakers of English (as speakers) and IL-English speakers<sup>36</sup> (as listeners), so that some understanding is reached of how Brazilians are understood when they signal important information by means of NS in interactions with members of the international community in general, not only with L1-English speakers.

In relation to the method for data collection used in these studies, production data is usually collected by means of reading (sentences or text passages) alone (Hahn, 2004; Manosso, 2013; Jenkins, 1997; Na et al., 2012) or with stimuli provided by the study researcher (Atechi, 2004; Tiffen, 1974). Perception/interpretability data is usually collected by means of listening to previously recorded data (Atechi, 2004; Hahn, 2004; Kivistö-de-Souza, 2017; Tiffen, 1974; Zoghbor, 2010). The only study that collected production and perception/interpretability data by means of face-to-face interactions between speaker and listener was Passarella-Reis and Silveira (2016), which reported the production data of the present research-study piloting. The scholars had the assumption that having speakers and listeners produce and hear speech in interactions makes this action of communicating (through controlled tasks in a laboratory) more similar to real life interactions than reading alone to a recording device and listening alone from an audio-player device, as speakers and listeners are more committed to conveying and to understanding the messages. The same data collection method is used in the present study.

With relation to the description of production data, studies usually describe the productions by displaying them in spectrograms figures only (de Castro Gomes, 2012; Harris, 2014), by describing the intonational patterns (Na et al., 2012), by using font emphasis in the orthographic description (Nava & Zubizarreta, 2010) or by mentioning the word where NS falls (Manosso, 2013). The only study that described production data by means of both spectrograms and font emphasis (with

---

<sup>36</sup> There is one L1-English participant. Based on the way English is understood in the present study, L1-English speakers are members of the international community as they also use English to communicate in international contexts.

some intonation description) was Passarella-Reis and Silveira (2016), which reported the production data of the present research study piloting. This way of describing production data is maintained in the present research study as it provides a clearer picture of the production when no audio is provided. Additionally, none of the studies mentioned above measured the pitch change range where NS falls, except for the latter. Measuring this range is important as it enables comparing productions among NS uses and participants. This measurement of pitch change range is also carried out in the analysis and description of production data in the present study.

Regarding the analysis of perception/interpretation data, most studies rely solely on listeners' answers on perception/interpretation sheets (Atechi, 2004; Jenkins, 1997; Hahn, 2004; Tiffen, 1974). The present study looks into the listeners' answers in the task sheets and their reports during the follow-up interviews. During the interviews, the video-recordings of the interactions for the tasks and the listeners' answers were considered so that researcher and listeners would discuss what contributed to the expected/unexpected perception of NS/interpretation of speakers' intent.

This chapter defined the most important concepts and reviewed the most relevant studies for the present study. The next chapter describes the method used.

### 3 METHOD

The present study investigated the sites of reduced interpretability caused by the unexpected allocation of nuclear stress by Brazilian speakers of English in face-to-face interactions with speakers of English from different L1 backgrounds. Thus, participants in this study were Brazilians and foreigners who were in Brazil at the time of data collection. This chapter is dedicated to explaining how these participants were recruited (Section 3.1), who they were (Section 3.2), the procedures and instruments of data collection (Section 3.3), and how data was analyzed (Section 3.4). The upcoming section describes how the participants were recruited and the criteria for selecting them.

#### 3.1 RECRUITING AND SELECTING PARTICIPANTS

Participants in the present study were 14 Brazilians and 14 members of the international community. In order to recruit participants, the researcher visited some English and Portuguese language classes offered by Extracurricular Courses in the Language Department of the Federal University of Santa Catarina (UFSC).

First, the coordinator of the Department was contacted in order to have his permission to have the research conducted with people attending classes offered by the Extracurricular Program.

Second, the coordinators of each area, English and Portuguese as a Foreign Language, were contacted and lists containing the schedule of teachers for levels 5, 6, and 7 in the case of English, and of levels 1, 2, and 3, in the case of Portuguese, were made available to enable the researcher to contact the teachers of those groups and get their permission to have their classes visited.

Third, the teachers were contacted via email and/or personal approach. In all, nine classrooms were visited: four groups of Portuguese for Foreigners (levels 1, 2, and 3) and five groups of English (levels 5, 6, and 7). Students were invited to participate and the ones willing to do so (N = 46) filled out a form with their names, contacts and availability (Appendices A and B).

Finally, all the volunteers (N = 23 foreigners and 23 Brazilians) were contacted via *Whatsapp*®, email and/or phone calls and only the ones whose availability conformed with the lab's schedule availability within the span of three weeks from that date were scheduled to

participate in the first phase of data collection (N = 16 foreigners and 15 Brazilians)<sup>37</sup>. The next section is dedicated to describing the first phase of data collection, which aimed at verifying the participants' level of oral proficiency in English.

### 3.1.1 Data collection for level of proficiency

As the present research is concerned with oral communication, the level of proficiency in English was an important variable for the study and thus had to be controlled. Speakers with a low level of proficiency in English were not desirable as participants due to the nature of the tasks involving data collection. For that reason, only students supposedly attending classes for longer than two years (Extracurricular Courses, level 5 on), in the case of Brazilians, were approached and invited to participate. For the foreigners, it was believed that they had at least an intermediate level as they would already use English to communicate internationally<sup>38</sup>. Despite these assumptions, an oral proficiency test was carried out in order to verify the participants' actual level of oral proficiency in English. Therefore, the purpose of the first phase of data collection was to collect oral data that enabled this verification.

Participants were 15 Brazilian Portuguese speakers of English and 16 speakers of English as an International language from different L1s. After recruiting the participants as described in Section 3.1, the data collection began.

They were contacted via *Whatsapp*®, email and/or phone calls in order to set the meeting for the first section of data collection. Each participant was sent a reminder message on the day before the meeting, which followed the steps described in the next paragraphs.

First, the participant was presented with, read, and signed the Consent Term (Appendices C and D), which consisted of information about the title of the study, its importance, the steps of data collection,

---

<sup>37</sup> Some volunteers never answered the *whatsapp*® message or email, others were not available at the times made available by the laboratory.

<sup>38</sup> The expectation that these participants would be able to communicate in English was mentioned when the invitation was made.

the risks and benefits, and the researcher's and the Ethics Board's contacts<sup>39</sup>.

Next, each participant filled out a questionnaire to gather information on personal data and their language experience (Appendix E). Additionally, follow-up questions were asked in order to make clear any obscure information that was therein provided. Foreigners were asked about the time they arrived in Brazil, their amount of exposure to spoken Portuguese, and their departure date. Brazilians were asked if they had attended public or regular schools and if they thought that school classes helped them learn English.

Subsequently, the participant was directed to the soundproof booth, where he or she received a detailed explanation on how to perform the oral task (adapted from Silveira, 2011; Appendix F). Besides the written instructions provided by means of slides, participants were told that they were free to speak as much as they wished, which could take them, for instance, from one to fifteen minutes. They were also told that they could skip the details they would not feel comfortable with. With the objective of offering some guidance during the description, they were presented with two questions: (1) What can you see in the picture? and (2) What is happening in the picture? The instructions provided an example in order to illustrate what the participant was expected to do. After testing the recording equipment<sup>40</sup>, the door was closed and the participant started the recording of his or her speech at the researcher's signal. Participants spoke from three to seventeen minutes. This was the last procedure of this phase of data collection, and thus, after doing the recording, the participant was dismissed. This data was collected from November 9 to November 12, 2015.

### **3.1.2 Data analysis for level of proficiency**

The audios were transferred to an audio editor software, were listened to, and a 25-to-30-second sample was extracted from each recording in order to generate the samples to be submitted to the raters. Each sample consisted of the description of part or of the entire picture

---

<sup>39</sup> The proposal for the present research was submitted to and approved by the Ethics Board of the Federal University of Santa Catarina (protocol number 1.323.812).

<sup>40</sup> The recording equipment is described in Table 10b, Section 3.3.

provided in the second slide of the slide show (Appendix F). The first slide was chosen not to be used because most participants reported being too anxious while performing the oral production task and the researcher thought the first slide would get more influenced by that feeling of anxiety than would the proceeding ones. Descriptions from the third slides on were not used because the third slide yielded enough material for the proficiency rating task.

Complete sentences were chosen to compose the samples because they “tend to be more informative and generate more adequate evaluations” (Silveira, 2011, p. 77, my translation)<sup>41</sup>. By using *Audacity*, a free audio editor and recorder, each sample was normalized at 60 dB and long pauses were removed. Two samples were randomly repeated in order to monitor the raters’ rating consistency. The final version of the speech file to be submitted to the raters to evaluate the proficiency level of the speakers then consisted of 33 25-to-30-second samples, separated by identifying numbers and silence gaps of three seconds, and had the total duration of 18 minutes and 20 seconds.

Eight highly proficient speakers of English, a Venezuelan and seven BP researchers in the area of Linguistics, participated as raters. In order to carry out the assessment of the informants’ level of oral proficiency, the raters were emailed an audio file for training, a text file containing instructions (Appendix G), a text file containing a form to perform the assessment itself (Appendix H), and a link to the audio file containing the speakers’ speech samples<sup>42</sup>. By following Silveira (2011), the raters were asked to make their assessment based on the “rater’s subjective notion of linguistic proficiency”, in order to “classify the informants according to their level of proficiency”<sup>43</sup> (Silveira, 2011, p. 82, my translation). The raters listened to the audio containing a sample of each informant and rated the informants’ level of oral proficiency by using a *Likert* scale from 0 (zero) – low oral proficiency level – to 10 (ten) – high oral proficiency level. After rating the participants, the raters emailed the file containing their ratings, which were transferred to an SPSS spreadsheet so that the statistical analysis could be performed.

---

<sup>41</sup> “tendem a ser mais informativas e a gerar avaliações mais adequadas” (Silveira, 2011, p. 77).

<sup>42</sup> This audio file was made available on a webpage for the rater to download because it was too heavy to be sent by email.

<sup>43</sup> “noção subjetiva de proficiência linguística do avaliador” ... “classificar os informantes de acordo com seu nível de proficiência” (Silveira, 2011, p. 82).



### 3.1.3 Results for level of proficiency

In order to verify the extent to which the raters agreed on their ratings, an inter-rater reliability test was run for looking at intraclass correlation, namely, Cronbach's Alpha. This test does "not only take into account the correlation between judges [raters], but also look[s] at whether the actual scores they gave participants differed" (Larson-Hall, 2010, p. 170). Tables 3 and 4 show the results of the Cronbach's Alpha and Table 4 displays the reliability analysis.

Table 3

*Cronbach's Alpha Output from the Reliability Analysis*

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Raters
.904	.920	8

Table 4

*Inter-rater correlation matrix – Reliability Analysis*

	Rater1	Rater2	Rater3	Rater4	Rater5	Rater6	Rater7	Rater8
Rater1	1.000	.462	.477	.502	.588	.501	.310	.533
Rater2	.462	1.000	.723	.407	.646	.763	.506	.680
Rater3	.477	.723	1.000	.743	.517	.753	.626	.648
Rater4	.502	.407	.743	1.000	.472	.679	.612	.623
Rater5	.588	.646	.517	.472	1.000	.612	.494	.631
Rater6	.501	.763	.753	.679	.612	1.000	.642	.693
Rater7	.310	.506	.626	.612	.494	.642	1.000	.661
Rater8	.533	.680	.648	.623	.631	.693	.661	1.000

Table 5  
*Rater-total statistics – Reliability Analysis*

	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Rater1	47.61	82.309	.595	.452	.900
Rater2	48.39	79.434	.728	.795	.891
Rater3	47.24	84.627	.803	.779	.895
Rater4	50.15	68.820	.702	.745	.895
Rater5	49.97	69.155	.692	.572	.896
Rater6	49.64	73.801	.830	.757	.880
Rater7	50.33	73.917	.683	.566	.893
Rater8	48.24	71.752	.804	.675	.881

Table 3 shows a Cronbach's Alpha of .904, which is considerably high. Table 4 shows the inter-rater correlation matrix. The paired correlations between raters were in the range of 0.31 to 0.76, which are medium and large effect sizes and, according to Larson-Hall (2010), can be fairly reliable. Table 5 displays the numbers if each rater individually were removed. The figures indicate that there was not too much variation (.88 to .90). This means that the variability in the scores was due to variation in the samples and not variation in the raters and thus, together with the results displayed in Tables 3 and 4, it indicates that the ratings given by these raters can be trusted. Table 6 provides the results of the participants' oral proficiency assessment.

Table 6  
*Oral Proficiency level of participants – speakers and listeners*

Level of proficiency	Brazilians	Rating	Foreigners	Rating
Not proficient				
Basic				
Intermediate				
Pre	B-01	4.63		
	B-02	4.63		
	B-03	4.76		
Inter			F-02	5.88
Upper	B-04	6.00	F-05	6.13
	B-05	6.13	F-16	6.38
	B-06	6.38	F-15	6.63
	B-07	6.38	F-03	6.63
	B-08	6.38		
	B-09	6.50		
	B-10	6.50		
	B-11	6.75		
Advanced	B-12	7.00	F-10	7.00
	B-13	7.25	F-06	7.00
	B-14	7.88	F-11	7.38
	B-15	8.00	F-09	7.50
			F-13	7.63
			F-01	7.88
			F-04	8.13
			F-07	8.50
L1-like			F-14	9.13
			F-08	9.50
			F-12	9.88

In order to fit each participant into a level of oral proficiency, the values in the 0-10 scale and oral proficiency levels respected the following correspondence:

- a) 0 – not proficient at all
- b) 1-3 – low oral proficiency (basic)
- c) 4-6 – medium oral proficiency (pre-intermediate: 4, intermediate: 5, upper-intermediate: 6)
- d) 7-8 – high oral proficiency (advanced)

e) 9-10 – high oral proficiency (L1-like)

The Brazilians' oral proficiency level varied from pre-intermediate up to advanced, with scores that ranged from 4.63 to 8.00, while that of the foreigners varied from intermediate to L1-like, with scores that ranged from 5.88 to 9.88. All in all, both Brazilians and foreigners were rated as having an intermediate or a higher level of oral proficiency and thus could actually participate in the study.

### 3.1.4 Summary

This section described how participants were recruited, the procedures for collecting data in order to have the participants' level of oral proficiency verified, and the analysis of these data. Although, in the beginning, 46 people volunteered, only 31 actually went through the procedures described above. Because results showed that their level of proficiency was intermediate or higher, all 31 volunteers were apt for actually participating in the next phase of the study.

Each participant was informed their level of proficiency by means of messages (*whatsapp*® or email) or phone calls and were invited to participate in the second phase of data collection, which is described in Section 3.3. The next section describes these participants.

## 3.2 PARTICIPANTS

Participants in the present were 14 Brazilians and 14 IL-English speakers. Initially, they were 15 L1-BP speakers of English (Brazilians) and 16 IL-English speakers (foreigners). For the reason that each Brazilian had to interact with an international speaker, the same number of foreigners and Brazilians was necessary. Therefore, one of the foreigners participants had to be dismissed, namely F-16. The criterion used to exclude that participant was schedule availability. Additionally, a Brazilian participant gave up on her participation and then the IL-English listener who would interact with her also had to be dismissed.

### 3.2.1 The Speakers

Since the objective of the present study was to understand how the intent of Brazilians was interpreted by the international community while communicating in English, speakers in the present study were originally 15 L1-BP speakers of English. Unfortunately, because a

member of one pair gave up on her participation due to schedule availability, the present study actually counted with the participation of 14 Brazilians only. For a quick look, Table 7 displays a summary of the data gathered about these participants and the following paragraphs are dedicated to their description.

Speakers were six males and eight females, who lived most of their lives in the South of Brazil (Blumenau: 1; Chapecó: 1; Criciúma: 1; Florianópolis: 7; Orleans: 1; São José: 1; and Turvo: 1), except for B-07, who lived most of her life in Rio. None of the speakers had lived abroad. All Brazilian participants were living in Florianópolis at the time of data collection for academic (N = 10), professional (N = 3) or personal reasons (N = 1). Their area of work/study were civil engineering (N = 3), business management (N = 1), Chemistry (N = 3), computer science (N = 1), dentistry (N = 1), design (N = 1), law (N = 1), and executive secretariat (N = 1); only speaker B-11 did not fit any category due to the fact he was still in high school. In fact, level of education varied among speakers. They were mostly undergraduate students (N = 10), one high school student, two holding undergraduate degrees and two specialists.

The speakers' ages ranged from 16 to 56 (M = 24.87). Regarding their experience with learning, they started having formal English classes at the age of seven (N = 4), 10 (N = 3), 11 (N = 4), 13 (N = 1), 15 (N = 1), and 18 (N = 1) (M = 10.57). Most speakers had English classes both in regular and language schools (N = 12) and two had English only in language schools, due to the fact they had Spanish (B-14) or French (B-12) at elementary and high school. At the time of data collection, they had been studying English for three to 12 years (M = 7.64). All of them were attending English classes at the time of data collection.

Table 7  
*Speakers' background data*

ID	Gender	Age	Education	Occupation	Lived	AOL	English classes	How long (year)	Prof. reported	Prof. verified	Speak (h/w)	Listen (h/w)	Write (h/w)	Read (h/w)
B-01	Female	56	Und. Degree	Retired (computer scientist)	Florianópolis	15	Regular/ language	6	Inter	Pre-inter	6	2	1	1
B-02	Female	25	Und. Student	Executive Secretariat	Florianópolis	11	Regular/ language	10	Pre-inter	Pre-Inter	1.5	18	2	1.5
B-03	Female	42	Specialist	Dentist	Florianópolis	10	Regular/ language	10	Pre-inter	Pre-Inter	3	4	0.5	0.5
B-05	Female	22	Und. Student	Civil Eng. Student	Criciúma	10	Regular/ language	9	Inter	Upper-inter	5	17	2	3
B-06	Male	19	Und. Student	Civil Eng. Student	Orleans	7	Regular/ language	10	Upper-inter	Upper-inter	3	7	1	4
B-07	Female	19	Und. Student	Law Student	Rio de Janeiro	7	Regular/ language	12	Inter	Upper-inter	3	5	1	1.5
B-08	Male	20	Und. Student	Engineering Student	Turvo	11	Regular/ language	8	Inter	Upper-inter	1	5	1	6
B-09	Female	22	Und. Student	Chemistry Student	São José	11	Regular/ language	11	Inter	Upper-inter	1	10	0.5	3
B-10	Male	27	Specialist	Business Manager (public server)	Florianópolis	7	Regular/ language	4	Inter	Upper-inter	3	6	3	6
B-11	Male	16	High school student	High School Student	Blumenau	7	Regular/ language	9	Inter	Upper-inter	10	20	5	3
B-12	Female	19	Und. Student	Chemistry Student	Florianópolis	13	Language <sup>c</sup>	6	Pre-inter	Adv.	1	8	1	6
B-13	Male	27	Und. Degree	Freelancer Designer	Florianópolis	10	Regular/ language	3	Inter	Adv.	2	6	1	6
B-14	Female	22	Und. Student	Secretariat Student	Florianópolis	18	Language <sup>c</sup>	4	Inter	Adv.	4	7	2	4
B-15	Male	18	Und. Student	Chemistry Student	Chapeçó	11	Regular/ language	5	Inter	Adv.	3	7	1	2

Lived: place participants lived most of their lives; AOL: age when participants started learning English/having English classes; English classes: where participants had English classes; How long: how many years participants have been studying/learning English; Prof. reported: level of oral proficiency participants reported; Prof. verified: level of oral proficiency as rated by the raters; Speak: how many hours per week participants speak English; Listen: how many hours per week participants listen to English; Write: how many hours per week participants write in English; Read: how many hours per week participants read in English.

The speakers' reported level of proficiency varied: pre-intermediate (N = 3), intermediate (N = 10), upper-intermediate (N = 1). The reported level did not match the verified level of proficiency in most cases: pre-intermediate (N = 3), upper-intermediate (N = 7) and advanced (N = 4). Speakers tended to report having a lower level of oral proficiency than that they actually had (N = 10). Three speakers' reported level matched their verified level and only one speaker reported having a higher level of proficiency.

Concerning speakers' use of English, amount of time varied among participants and skills. Speakers would speak English from one to 10 hours per week (M = 3.32 hours), listen to English from two to 20 hours per week (M = 8.71 hours), write in English from half an hour to five hours per week (M = 1.57 hours), and read English texts from half an hour to six hours per week (M = 3.39 hours). Most speaking was carried out in English classes, except for participants B-11 and B-01. Out of the four skills, listening to English was the most practiced skill by the speakers. This leads us to a mismatch between the skills linked to oral communication, namely speaking and listening, since these participants would spend more time listening to than speaking English.

### **3.2.2 The Listeners**

Given the objective of the present study was to examine how Brazilians are understood by the international community when they communicate in English, listeners were 15 members of the international community who could speak English. Because a speaker gave up on her participation, as mentioned in the previous section, one listener had to be dismissed due to the design of data collection, which takes place by means of pairing together a BP-L1 speaker of English and a listener who is a speaker of English from a nationality other than Brazilian. Hence, F-04 was dismissed. I would like to clarify that there was no interest in listeners from specific countries and that the countries represented here were not chosen from a great variety of countries available; they were the international community members available and willing to participate at the time of data collection. For a quick look, Table 8 displays a summary of the data gathered from these participants and the following paragraphs are dedicated to their description.

Table 8

*Listeners' background data*

ID	Gender	Age	Nation ality	Education	Occupation	Lived O L	A O L	How long (year)	Prof. reported	Prof. verified	Speak (h/w)	Listen (h/w)	Write (h/w)	Read (h/w)	In BR	Liste n to BP
F-01	Female	25	German	Und. Student	Und. Student	Germany	11	13	Upper-inter	Adv.	30	6	1	1	4	7
F-02	Male	22	French	Und. Student	Und. Student	France	11	11	Upper-inter	Inter	0	20	0	0	3	56
F-03	Male	36	Italian	PhD. Degree	Professor	Italy	14	5	Upper-inter	Upper- Inter	4	4	10	10	3	8
F-05	Male	22	Polish	Und. Degree	Und. Student	Poland	10	12	Upper-inter	Upper- Inter	5	4	3	7	7	3
F-06	Female	35	Italian	PhD. Degree	In between jobs	Italy	9	10	Inter	Upper- Inter	2	2	2	2	3	14
F-07	Male	25	Norwegian	Und. Student	Und. Student	Norway	8	10	Adv.	Adv.	5	10	2	5	4	49
F-08	Female	21	German	Und. Student	Und. Student	Germany	9	10	L1-like	L1-like	7	7	2	2	3	56
F-09	Male	27	French	PhD. Degree	Post Doc Researcher	France	11	11	Upper-inter	Adv.	2	2	0.5	0.5	3	28
F-10	Female	33	Chilean	Und. Degree	Graphic Designer	Chile	4	14	Adv.	Adv.	0	28	0	15	12	56
F-11	Female	23	Slovak	Und. Degree	Mother	Slovakia	7	10	Upper-inter	Adv.	70	45	14	5	16	70
F-12	Male	23	English	Master Degree	Chemical Engineer	England	0	0	L1-like	L1-like	27.9	27.9	9.3	18.3	42	1.5
F-13	Male	28	Russian	Und. Degree	Mobile Developer	Belarus	11	17	Pre-inter	Adv.	16	26	10	35	12	56
F-14	Male	24	Swiss	PhD. Student	PhD Student	Switzerland	13	5	L1-like	L1-like	6	5	0.5	1	4	56
F-15	Female	21	Japanese	Und. Student	Und. Student	Japan	4	6	Inter	Upper- Inter	5	7	1	1	3	63

Lived: place participants lived most of their lives; AOL: age when participants started learning English/having English classes; How long: how many years participants have been studying/learning English; Prof. reported: level of oral proficiency participants reported; Prof. verified: level of oral proficiency as rated by the raters; Speak: how many hours per week participants speak English; Listen: how many hours per week participants listen to English; Write: how many hours per week participants write in English; Read: how many hours per week participants read in English; In BR: how long (measured in months) participants have been in Brazil; Listen to BP: how many hours per week participants listen to Brazilian Portuguese.



Listeners were eight males and six females, who lived most of their lives in their countries of origin (England: 1; France: 2; Germany: 2; Italy: 2; Japan: 1; Norway: 1; Poland: 1; Russia: 1; Slovakia: 1; Chile: 1; and Switzerland: 1). All listeners were living in Florianópolis at the time of data collection for academic (N = 8), professional (N = 5) or personal (N = 1) reasons. They were from different areas of study/work: computer science (N = 2), architecture (N = 1), business engineering (N = 1), chemistry (N = 3), electrical engineering (N = 1), electronics engineering (N = 1), graphic design (N = 1), marketing management (N = 1), mechanical engineering (N = 1), physical education (N = 1), and physics (N = 1).

Regarding their level of education, listeners were five undergraduate students, four people holding undergraduate degrees, a master student, a person holding a master degree, and three people holding PhD degrees. At the time of data collection, their occupations were: undergraduate student (N = 6), PhD student (N = 1), chemical engineer (N = 1), graphic designer (N = 1), mobile developer (N = 1), professor (N = 1), post doc researcher (N = 1), and mother (N = 1). One of the participants was between jobs.

The listeners' ages ranged from 21 to 36 (M = 26.33). Except for the L1-British-English listener, listeners started having formal English classes at the age of four (N = 2), seven (N = 1), eight (N = 1), nine (N = 2), 10 (N = 1), 11 (N = 4), 13 (N = 1), and 14 (N = 1) (M = 8.71). Listeners had English classes in their regular schools (N = 14) during a period of five to 17 years (M = 9.57). None of the listeners was attending English classes at the time of data collection.

The listeners' reported level of proficiency also varied: pre-intermediate (N = 1), intermediate (N = 2), upper-intermediate (N = 6), advanced (N = 2), and L1-like (N = 3). Their verified level of proficiency showed they were intermediate (N = 1), upper-intermediate (N = 4), advanced (N = 6), and L1-like (N = 3). Half of them had their reported level of proficiency matching their verified level of proficiency, while the other half would report a level that was lower (N = 6) or higher (N = 1) than the verified level.

Concerning listeners' use of English, amount of time varied among participants and skills. Two listeners reported not speaking English at all at the time of data collection (F-02 and F-10). The remainder of the listeners would speak English from two to 70 hours per week (M = 12.85 hours). Listeners reported they would listen to English

from two to 45 hours per week ( $M = 13.85$  hours), write in English from half an hour to 14 hours per week, except for two participants who would not write in English at all, namely F-02 and F-10 ( $M = 3.95$  hours), and read English texts from half an hour to 35 hours per week, except for one participant who would not read in English at all, namely F-02 ( $M = 7.34$  hours). By considering the mean time of English use for speaking and listening, skills used in oral communication, one can suggest that, in general terms, these listeners would spend nearly the same amount of time speaking to and listening to their interlocutors during their interactions in English.

These listeners had been living in Brazil for different amounts of time. Most of them were exchange students that had been in Brazil for a short time: three months ( $N = 6$ ) and four months ( $N = 3$ ). Two other exchange students had been in Brazil for a little longer: seven months and 12 months, while other two were in Brazil for professional reasons for 16 months and 42 months. The mean time in Brazil was 8.5 months at the time of data collection.

Regarding their contact with BP language, all of them were attending BP language classes at the time of data collection and would spend from one and a half to 70 hours listening to BP during classes or during interactions with Brazilian people<sup>44</sup> ( $M = 37.39$  hours).

### 3.2.3 Summary

Section 3.2 described the participants of the present study: 14 L1-BP speakers of English as speakers and 14 speakers of English from a nationality other than Brazilian as listeners. The mean age of speakers and listeners was somewhat alike: 24.87 and 26.33 years, respectively. While all listeners learned their English firstly from their regular curriculum at school and later from international interaction, most Brazilians had to enroll in language courses in order to complement their learning. In general, speakers started learning English at a later age than did the listeners (10.57 and 7.64 years, respectively) and studied English for a shorter period of time than did the listeners (8.71 and 9.57, respectively). While most speakers had an intermediate level of oral proficiency, most listeners had an advanced level of oral proficiency.

---

<sup>44</sup> These members of the international community interacted with Brazilians sometimes in English and eventually in BP, depending on their interlocutors' level of proficiency in English. When one language would not work, they would use the other.

This difference is also reflected in their English use: The mean time speaking English was quite shorter for the speakers than it was for the listeners (3.32 and 12.85 hours, respectively). Speakers and listeners were very willing to engage in the activities and steps of data collection, which are described in the next section.

### 3.3 DATA COLLECTION

#### 3.3.1 Instruments and procedures

As the aim of the present study was to investigate the interpretation of L1-BP speakers' intent based on their use of NS in English utterances in interactions held with other speakers of English from different nationalities, participants in this study were Brazilians and speakers of English from nationalities other than Brazilian. In order to have participants to interact, they were divided into 14 pairs composed by a Brazilian and a foreigner<sup>45</sup>. One out of 16 foreigners who volunteered to participate was chosen not to participate in the present study, namely F-16. The criterion used to leave this participant aside was schedule availability.

In the interactions, the Brazilians acted as speakers and the foreigners acted as listeners. They were contacted through messages (*whatsapp*® or email) or phone calls in order to find the day and the time that would be fine with their schedules and that of the Language Laboratory, where data were collected, so that participants could meet. Each participant was paired with one another based on the criterion of roles being played (speaker vs listener) and schedule availability. Table 9 shows each pairing made.

---

<sup>45</sup> Recall that this study initially had 15 pairs, but one pair could not proceed with data collection because a member of the pair gave up on her participation, as mentioned earlier in Section 3.2.

Table 9  
*Pairs for the interactions for data collection*

Pairs (N = 14)	Brazilian Speakers (N = 14)	Foreigner Listeners (N = 14)	Level of Proficiency
1	B-01	F-01	Pre X Advanced
2	B-02	F-02	Pre X Inter
3	B-03	F-03	Pre X Upper
4	B-05	F-05	Upper X Upper
5	B-06	F-06	Upper X Advanced
6	B-07	F-07	Upper X Advanced
7	B-08	F-08	Upper X L1-like
8	B-09	F-09	Upper X Advanced
9	B-10	F-10	Upper X Advanced
10	B-11	F-11	Upper X Advanced
11	B-12	F-12	Advanced X L1-like
12	B-13	F-13	Advanced X Advanced
13	B-14	F-14	Advanced X L1-like
14	B-15	F-15	Advanced X Upper

Each one of the pairs (Table 9) met the researcher at the laboratory. The members never arrived together and were unknown to one another. At the time a member of a pair arrived, he or she was told that his or her partner in data collection would not be introduced to him or her, since they were supposed to hold a first meeting conversation by their own as part of data collection. In order to minimize the discomfort caused by the waiting due to occasional delays, the researcher would hold an informal conversation with the participant arriving first, talking mainly about their place of origin, by looking information on the Internet about the city (population, climate, location and the like).

As soon as the second member of the pair arrived, data collection actually began. The speaker (Brazilian) and the listener (foreigner) were introduced to the context of data collection. They were told that they both worked in an international company settled in Canada and that they had coworkers from different nationalities and that their contact language was English. Although they worked at the same company, they had never met before and had to get to know each other. They had some friends in common, namely Anna, Maria, John, and Zenny, the latter being the most wonderful person they were aware of. Zenny was always there for them. If they had problems, Zenny would have the solution.

Zenny was always a helping hand. This short introduction had the aim of involving participants and helping them connect with each other in order to hold a meaningful and purposeful interaction. Table 10a summarizes the tasks, the roles of listeners and speakers and includes the Appendices related to each task.

After setting the context of data collection, participants were given their first task: *Breaking the Ice* (Appendix I). For Task 1, they entered the soundproof booth. The speakers were both audio and video recorded, while the listeners were only audio recorded in Task 1. Table 10b describes the equipment used during the tasks and previously when data for proficiency level rating were collected.

Task 1 was inspired in Trofimovich and Kennedy (2014). In this task, participants had to get to know each other, by asking for information they would naturally ask someone in a first meeting, such as name and place of origin. Additionally, they were asked to find three things they had in common, for instance, a dislike of sushi (Trofimovich & Kennedy, 2014). The purpose of this task was to break the ice and help them feel more comfortable in the presence of each other<sup>46</sup>.

After Task 1, participants got out of the soundproof booth and were told that the Company was facing hard financial times and, because of that, it was planning to fire some of the employees. The Company had a tentative firing list and the listener's name was on it. The decision about whether to fire or not the listener would depend on the listener's performance in Tasks 2, 3, and 4. The listener was told that his position at the Company required him or her to listen well and understand people from different parts of the world. The forthcoming tasks would require him to listen to the speaker and try to grasp the speaker's intention in each utterance said. The speaker, in turn, was told that the speaker was responsible for helping the listener keep the listener's job by speaking as intelligibly as possible. Telling participants that the listener's job was at risk had twofold effects. On the one hand, the speaker tried harder to speak clearly and convey the message imposed by the discursive context in order to help the listener succeed in his task of grasping the speaker's intentions. On the other hand, the listener himself/herself struggled to understand what had been said by the speaker and to get the speaker's intentions in order to keep his or her

---

<sup>46</sup> Although it was not the purposes of the present study, the data collected by means of this task can be analyzed in future studies in order to look into NS placements by both Brazilians and members of the international community.

job, which made the tasks purposeful and led both listener and speaker to fully engage in the pairwork activity.

Table 10a

*Tasks, roles of speakers and listeners, and respective Appendices for the tasks*

Tasks	Roles		Appendices
	Speaker	Listener	
Task 1 - Breaking the Ice	Listener's coworker in an International Company. Had to get to know the listener and find three things they had in common.	Speaker's coworker in an International Company. Had to get to know the speaker and find three things they had in common.	I
Task 2 – I know the answer!/I don't know the answer, do you?	Help the listener keep the listener's job by speaking as intelligibly as possible. Analyze the context question for each utterance and read aloud the answer with the NS on the information being elicited.	One of the names in the company's firing list. Had to show listening skills in order not to be fired. Had to understand what piece of information was being elicited and check one of the three options presented in the listener's sheet.	J K
Task 3 – You're wrong...	Analyze the context utterance and read aloud the response-utterance with the NS on the information being corrected.	Had to understand what piece of information was being corrected and check one of the three options presented in the listener's sheet.	L M
Task 4 – I know what you're saying next!	Analyze the context disjunctive question and read aloud its first clause with the NS on the information being contrasted.	Had to guess what the second clause of the speaker's question would be, by identifying the first element of the contrast made in the first clause and then choose one of the three options presented in the listener's sheet.	N O

Table 10b  
*Equipment used during data collection*

Data collection	Participants	Equipment
Oral Proficiency Task	Speakers and listeners	Audio recording: a <i>C 520 L</i> professional head-worn condenser microphone, connected to <i>MOTU Ultra Lite mk3</i> , a hybrid audio interface, and <i>OcenAudio</i> (2015), audio editor software <sup>47</sup> .
Tasks 1, 2, 3, and 4	Speakers	Audio recording: a <i>C 520 L</i> professional head-worn condenser microphone, connected to <i>MOTU Ultra Lite mk3</i> , and <i>OcenAudio</i> (2015). Video recording: a Nikon camera full HD.
Task 1	Listeners	Audio recording: a <i>Shure</i> microphone connected to a Sony voice recorder.
Interview with the listeners	Listeners	Audio recording: a <i>C 520 L</i> professional head-worn condenser microphone, connected to <i>MOTU Ultra Lite mk3</i> , and <i>OcenAudio</i> (2015). A laptop computer <i>Dell Vostro 5470</i> to play the videos recorded in the tasks.

In Tasks 2, 3, and 4, the Brazilians played the role of speakers and the foreigners played the role of listeners. The speakers were audio and video recorded. These procedures generated the samples to be analyzed in order to investigate the assignment of NS by Brazilian speakers of English. The videos were transferred to a laptop computer so that they could be assessed during the interview with the listener (foreigner), as described later in this section. Additionally, the videos were assessed during data analysis as they provided visual input of the speakers, such as rapid eyebrow movements, which were considered in

<sup>47</sup> *OcenAudio* was developed by LINSE – Laboratory of Circuits and Signal Processing – a research group at Universidade Federal de Santa Catarina. It is available at <http://www.ocenaudio.com.br/>.

order to enlighten information about the expected interpretation of unexpected productions<sup>48</sup>.

Having the context established, both speaker and listener were introduced to one of the three interpretability activities that comprised data collection, namely, Task 2: *I know the answer!/I don't know the answer, do you?* (Appendices J and K). In general terms, in the three interpretability tasks, listeners had to listen to the speakers and identify the speakers' intentions (Nelson, 2011). Even though in Task 1 speakers and listeners could speak freely in order to accomplish the task, the interpretability tasks involved reading.

Reading is one of the most used methods in the history of pronunciation research due to its manifold advantages: control for features and other elements such as vocabulary choice or grammar usage (Levis, 2011). However, as Levis highlights, reading aloud is a reading skill not a speaking skill, it promotes different performances (better or worse than in free speech) depending on the speaker who reads/speaks, and there is an agreement that individuals feel weird when reading aloud because it is not a common activity. However, as the focus of the present study is on interpretability linked to pronunciation rather than on overall intelligibility (linked to choice of words and grammar usage, for example), the only way of trying to avoid these intervening variables was by using more controlled tasks, made possible through reading aloud activities. Although reading aloud is sometimes only linked to decoding speech, it is important to highlight that, for the purposes of the present study, the speaker had to read and interpret the utterances in order to convey a specific message. Additionally, having a listener with whom to interact made the reading purposeful. In the pilot study, this was clear given that participants struggled in order to make their speech clear and showed to be concerned with the listeners' comprehension of the message. This concern with communicating brought the reading aloud act closer to really communicating a message in free speech. In order to have a greater picture of the pilot study, refer to Passarella-Reis and Silveira (2016).

Each of the three tasks required the speakers to communicate their intentions in different manners. Task 2 required them to answer to

---

<sup>48</sup> One of the limitations of the pilot study was the fact that the speakers were only audio recorded. Research has shown that visual cues such as rapid eyebrow movements may play a role during the production of nuclear stress (Krahmer & Swerts, 2007). Therefore, besides being audio recorded, speakers were also video recorded in the present study.



questions, focusing on the information supposedly being sought. Task 3 required them to correct information, by focusing on the information being corrected. Finally, Task 4 required them to contrast information in questions, by assigning a relevant pitch change to one of the elements in contrast in the first clause of disjunctive questions. Therefore, speakers (Brazilian speakers of English) produced samples in which they were answering questions, correcting mistaken information, and contrasting information (by means of NS), which are part of one's daily communication acts. The design for Tasks 2 and 3 were partially inspired by instruments used in Tiffen (1974) and Atechi (1994), while the design for Task 4 was partially inspired by the study conducted by Jenkins (1997).

Each of the three tasks (2, 3, and 4) was comprised by nine context sentences and questions (in the case of Tasks 2 and 4) or only context sentences (in the case of Task 3). These context inputs triggered the production of nine utterances, which had the following variables controlled: (1) NS position (initial, medial, or final) and (2) complexity of words (non-challenging, short challenging, long challenging) in the remainder of the sentence<sup>49</sup>. The utterances for each task are provided in the examples below, from (28) to (54)<sup>50</sup>.

#### Task 2

- (28) ZENNY loves you.
- (29) ZENNY sent you the **pearls**.
- (30) ZENNY **traveled** to **Pennsylvania**.
- (31) Zenny LOVES you.
- (32) The boss LOVES your **world**.
- (33) Zenny DISLIKES **vinegar**.
- (34) Ana loves ZENNY.
- (35) The **government** talked to ZENNY.
- (36) The T-shirt **shrank** A LOT.

---

<sup>49</sup> The pilot study results showed that certain words were difficult to pronounce and interfered with the way portions were highlighted in the utterances. Therefore, I decided to further investigate how much the presence or absence of complex words would affect NS production and perception.

<sup>50</sup> Words written in CAPITAL LETTERS are nuclear stressed, while **boldfaced** words are challenging words in terms of pronunciation.

## Task 3

- (37) No, ANA sent the e-mail.
- (38) No, MARIA had it **curled**.
- (39) No, RED is an **appetite** stimulator.
- (40) No, Ana LOVES dogs.
- (41) No, **pearls** MELT in **vinegar**.
- (42) No, yellow STIMULATES **optimism**.
- (43) No, Zenny loves DOGS.
- (44) No, the **rhythm suits** ZENNY.
- (45) No, **unfortunately** Zenny AGREED.

## Task 4

- (46) Do MEN lie a lot...?
- (47) Is ZENNY afraid of **heights**...?
- (48) Is ZENNY leaving **immediately**...?
- (49) Does the meeting START at ten...?
- (50) Does Zenny LIKE **shrimp**...?
- (51) Does this sequence REPRESENT the **hierarchy**...?
- (52) Does the party start at TEN...?
- (53) Did Ana **twirl** HER HAIR...?
- (54) Is the **atmosphere** around Zenny GOOD...?

Along the tasks, speakers had to read utterances with non-challenging words and utterances with two types of challenging words: short words and long words. The choice for the inclusion of challenging words was made because one of the objectives of the present study was to systematically investigate the extent to which the presence of challenging words in terms of pronunciation would affect accurate NS placements. Challenging words were chosen by chance as the discursive contexts and utterances were being devised for the purposes of the present study. They only had to meet the following criteria in order to be considered challenging:

- a) Short-challenging words needed to have one or two syllables (e.g., *twirl*, *shrimp*, *world*, *rhythm*) and usually have consonant clusters (e.g., *shrank*).
- b) Long-challenging words needed to have three or more syllables (e.g., *hierarchy*, *optimism*, *unfortunately*).

Non-challenging words varied in number of syllables and had a rather simple syllable structure (e.g., *loves*, *ten*, *meeting*) or were very frequent words (e.g., *start*).

In the three tasks, NS location also varied, given that investigating the effect of expected NS position in the utterance was also one of the objectives of the present study. In (28), (29), (30), (37), (38), (39), (46), (47) and (48), NS (in capital letters) is in initial position. (28), (37), and (46) include no apparently challenging words to L1-BP speakers of English, while (29), (30), (38), (39), (47), and (48) include at least one challenging word<sup>51</sup> (in boldface letters). For instance, the word *immediately* in (48) is a long word that showed to be challenging for participants to produce in the pilot study, who often misplaced word stress and thus changed some vowel sounds. Likewise, the word *pearls* in (29) is a short word that also showed to be challenging for participants in the pilot study given the final consonant cluster. Other words that had the potential of being troublesome in terms of pronunciation in these utterances were *travelled*, *Pennsylvania*, *curled*, *appetite*, and *heights*.

In (31), (32), (33), (40), (41), (42), (49), (50), and (51), NS is in medial position. (32), (33), (41), (42), (50), and (51) include possibly challenging words (*world*, *vinegar*, *pearls*, *optimism*, *shrimp*, and *hierarchy*, respectively) in initial or final position, while (31), (40), and (49) are expected to pose no obstacles in terms of pronunciation.

In (34), (35), (36), (43), (44), (45), (52), (53), and (54), NS falls in final position. (34), (43), and (52) include no apparently challenging words to L1-BP speakers of English, while (35), (36), (44), (45), (53), and (54) include the words *government*, *shrank*, *rhythm*, *suits*, *unfortunately*, *twirl*, and *atmosphere*, which are probably challenging words for these speakers as regards pronunciation.

The sequence of the utterances in each task was randomized. Additionally, the first sentence was repeated in the last position in each set, so that the results were not influenced by the inexperience of the speaker and the listener with the task. Overall, 378 utterances followed by repetitions when required (N = 123) were produced for analysis.

For collecting these data, directions were given to participants separately. For Task 2 (Appendix J), the speaker was explained that his

---

<sup>51</sup> It should be noted that some utterances include more than one challenging word – (30), (41), and (44) – and two words were repeated – *vinegar* and *pearls*. It was not planned, but these words fit well the contexts being created for the utterances and were then used.

or her coworker (the listener) had previously asked the speaker some questions, whose answers were unknown to the speaker, who promptly offered to go and try to find the answers to those questions. After some search on the net and few interviews around the company, the speaker managed to find all the answers and was back to provide the listener with them. So, for Task 2, the speaker would have to read silently a context sentence and a context question to understand the discursive act, and then read aloud the statement with the information which was supposedly being sought by the listener. An example is provided in (55).

(55) Speaker

Remember that you work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English. For this task, you will consider different context questions, by reading them silently. Each question establishes a context that requests you to say a piece of information that your coworker needs. You will say aloud that piece of information. To illustrate, see an example below.

(Talking to a coworker, who is a great fan of Adele.)

Read it silently: *When will Adele be in Canada?*

Say it aloud: **Adele will be in Canada in December.**

In the example above, the speaker needed to have in mind that the question being answered was *When will Adele be in Canada?* as if the question had been asked by his or her coworker (the listener). When the speaker read the answer, the speaker had to emphasize *December*, because that was the information the listener was looking for. When told to emphasize the information, the speakers asked how they were supposed to emphasize that piece of information and were informed that they were expected to find their own way of doing so. This reaction to the task signaled that they were not aware of the use of NS in English, which was confirmed by an informal interview by the end of this data collection session. The speaker was told not to repeat an answer unless the listener asked the speaker to. Additionally, the speaker was oriented to pause after each production in order to give some time for the listener to process the information. Finally, the speaker was advised to try to speak in the speakers' natural speed and to avoid breaking up the pieces

of information. These three last orientations (about repetition, pauses, and speed of speech) were also given before Tasks 3 and 4.

For Task 2 (Appendix K), the listener was told that the speaker (his or her coworker) was going to answer some questions the listener had previously asked. In his sheet, the listener was supposed to choose, out of a list of three questions, the one the speaker would be answering, according to the emphasis given to the information. An example is provided in (56).

(56) Listener

Remember that you work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English. For this task, you will consider that you have previously asked your coworker some questions. By the time you asked the questions, your coworker didn't know the answers. However, after a little time, your coworker came out with the answers. Your job is to listen to the answers and figure out what question he or she is answering. To illustrate, let us consider the following answer.

**Adele will be in Canada IN DECEMBER.**

The words in capital letters indicate that IN DECEMBER is emphasized. This emphasis means that IN DECEMBER is the most important information in the sentence and it is the information you were looking for. In your sheet, you will have the questions you have previously asked your coworker, as illustrated below.

**(A) Who will be in Canada in December?**

**(B) Where will Adele be in December?**

**(C) When will Adele be in Canada?**

In (56), the listener needed to have in mind that, by highlighting *December*, the speaker's intention had been to inform neither where (in Canada) nor who (Adele), but rather when (in December) Adele would be in Canada, and thus the listener should choose alternative (C). In case the listener did not get to understand which question was being answered after listening to the speaker's production of an answer, the listener was told that s/he ask only twice for the speaker to repeat a

given answer. If after the repetitions, the listener could not make an informed decision, the listener was asked to guess. The listener was expected to follow these pieces of advice in Tasks 3 and 4 too.

After Task 2, participants had a short break out of the booth, while they individually received information about Task 3 (Appendices L and M: *You're wrong...*). The speaker was explained that his or her coworker (the listener) was usually right about things. However, for this once, the listener was mistaken about some pieces of information. The speaker's task was to correct those pieces of information. In order to understand what piece of information to correct, the speaker had to read silently a situational sentence and a context sentence to understand the discursive act, and then read aloud the statement with the information being corrected. An example is provided in (57).

(57) Speaker

For this task, you will consider different sentences, by reading them silently. Each sentence establishes a context that requests you to correct a piece of information given by your coworker. You will read aloud the utterance that corrects the information. To illustrate, let us see an example below.

(Talking about your coworker John.)

Read it silently: John bought a car.

Say it aloud: **No, John bought a bike.**

In (57), the speaker needed to have in mind that the mistaken information was the item purchased by John, and thus should read the correction and assign more prominence to the word *bike*.

For Task 3, the listener was told that the listener had some mistaken pieces of information, which were going to be corrected by the speaker. In his/her sheet, the listener was supposed to choose, out of a list of three pairs of contrasting pieces of information, the contrasting pieces meant by the speaker, according to emphasis assignment. An example is provided in (58).

(58) Listener

You are a very-updated person. During your coffee break, you are having a conversation with your coworker about different subjects. Even though you are that updated, some information that you provide seems to be mistaken and your coworker corrects you. For this task, you will hear the information from your coworker and identify the piece being corrected.

To illustrate, consider that your coworker say:

**No, John bought a BIKE.**

The words in capital letters indicate that BIKE is emphasized. This emphasis means that BIKE is the information replacing the mistaken information provided by you. In your sheet, you will have possible corrections, as illustrated below:

**(A) John did, not Peter.**

**(B) John bought it, he did not sell it.**

**(C) A bike, not a car.**

In (58), the listener needed to have in mind that, by emphasizing the word BIKE, the speaker was correcting neither the agent (John) nor the action (bought), but rather the item purchased (a bike), and thus the listener had to choose the contrasting pair in (C): *A bike, not a car*.

After Task 3, participants had another short break out of the booth, while they individually received information about Task 4 (Appendices N and O: *I know what you're saying next!*). The speaker was explained that his or her coworker (the listener) liked finishing the listener's interlocutor's sentences. The speaker had many doubts about varied aspects in the company and knew that the listener could help the speaker with some of them. Therefore, the speaker wanted to ask the listener some questions. Nevertheless, the speaker was not supposed to finish the speaker's questions, that is, the speaker had to interrupt the question as if the listener were interrupting the speaker. In order to understand the context, the speaker had to read silently a situational sentence and the entire question to be asked, and then read aloud only the first half of the question. An example is provided in (59).

(59) Speaker

For this task, you will consider different questions, by reading them silently. Each question represents a whole question that you want to ask. Note, however, that your coworker likes finishing his/her interlocutors' questions and will not allow you to finish yours. To illustrate, let us consider the example below.

(Talking about Maria's trip.)

Read it silently: Is Maria traveling to Canada tomorrow or to England?

Ask it aloud: **Is Maria traveling to Canada tomorrow ...?**

In (59), the speaker should consider both the context sentence (*Talking about Maria's trip.*) and the question meant to be asked (*Is Maria traveling to Canada tomorrow or to England?*), by reading them silently. The question signals that the speaker is in doubt about the destination (Canada or England). Therefore, when the speaker actually asked the question, he or she emphasized the word *Canada* in order to signal that the unsaid second half of the question should contain another destination.

For Task 4, the listener was told that the listener was a very sensitive and smart person, who was very fast at getting what people say and that the listener loved to finish the listener's interlocutor's lines. The listener was informed that the speaker was going to ask the listener many questions, but that the listener would hear only the first half of the questions because the speaker would not finish them, as if the speaker were interrupted by the listener. The listener had, then, to figure out the missing second half, by choosing one out of the three finishing lines in his/her sheet. An example is provided in (60).

(60) Listener

You are very sensitive and smart. You are fast at getting what people are saying and you love to finish your interlocutors' lines. For this task, your coworker will ask you many questions, but you will hear only the first half of the questions as if you were interrupting him/her. Your job is to figure out what the second half would be, based on the way your coworker says it. To



illustrate, let us consider that your coworker ask the question below.

**Is Maria traveling to CANADA tomorrow ...?**

The words in capital letters indicate that CANADA is emphasized. Thus, your coworker knows that Maria is traveling tomorrow but he/she is in doubt about the destination. In your sheet, you will have three possible second halves, as illustrated below.

**(A) or is it John?**

**(B) or to England?**

**(C) or the day after tomorrow?**

In (60), the listener had to check alternative (B), as the speaker was in doubt about neither the agent (Maria or John) nor the day (tomorrow or the day after tomorrow), but rather the destination (Canada or England), by emphasizing the word *Canada*.

So as to avoid that the sequence of the tasks would affect speakers' and listeners' performances, the sequence of Tasks 2, 3, and 4 was altered, as illustrated in Table 11.

In the way it was organized, each five pairs (except for Sequence 3, which had four pairs) followed one of the sequences established for the tasks. The duration of each task varied among the pairs. The duration for Task 2 ranged from 2 min 12 s to 5 min 11 s. The duration for Task 3, in turn, ranged from 1 min 41 s to 4 min 39 s. Finally, the duration for Task 4 ranged from 2 min 18 s to 4 min 56 s.

Table 11  
*Sequences of Tasks 2, 3, and 4.*

	Sequence of the tasks (N = 3)	Pairs (N = 15)
Sequence 1	Task 2 – <i>I know the answer!</i>	BP-01 and IL-01
	Task 3 – <i>You're wrong!</i>	BP-05 and IL-05
	Task 4 – <i>I know what you are saying next</i>	BP-10 and IL-10
		BP-11 and IL-11
Sequence 2	Task 3 – <i>You're wrong!</i>	BP-12 and IL-12
	Task 4 – <i>I know what you are saying next</i>	BP-02 and IL-02
	Task 2 – <i>I know the answer!</i>	BP-07 and IL-07
		BP-08 and IL-08
Sequence 3	Task 2 – <i>I know the answer!</i>	BP-09 and IL-09
	Task 4 – <i>I know what you are saying next</i>	BP-13 and IL-13
	Task 4 – <i>I know what you are saying next</i>	BP-03 and IL-03
	Task 2 – <i>I know the answer!</i>	BP-04 and IL-04 <sup>a</sup>
	Task 3 – <i>You're wrong!</i>	BP-06 and IL-06
	BP-14 and IL-14	
	BP-15 and IL-15	

<sup>a</sup>. Data not collected as planned, given that the speaker quit the study.

After finishing the tasks, the participants took a *Word Familiarity Test* (Appendix P) and a *Pronunciation Self-Evaluation Test* (Appendix Q). In both tests, participants had to rate, in a scale from 01 (one) to 09 (nine), how acquainted they were with or how well they pronounced the words that were included in Tasks 2, 3, and 4. The objective of the *Word Familiarity Test* was to assess how acquainted the participants were with the vocabulary included in Tasks 2, 3, and 4, whereas the *Pronunciation Self-Evaluation Test* investigated how well they thought they pronounced those words in the tasks. In the *Word Familiarity Test*, 01 (one) meant that the participant did not know the word at all, while 09 (nine) meant that the participant was very acquainted with the word. In the *Pronunciation Self-Evaluation Test*, 01 (one) meant that the participant thought that the participant's pronunciation of that word was not intelligible at all, while 09 (nine) meant that the participant thought that the participant's pronunciation of the word was very intelligible. The *Word Familiarity Test* was used as one more resource to help understanding the reasons why the nuclear stress was not placed as expected when it was the case and, when it was, the reasons why it was not interpreted as expected. The *Pronunciation Self-Evaluation Test* helped understand the way both speaker and listener perceived their

pronunciation. For the speakers, having to produce troublesome words in terms of pronunciation could be a distractor. The results of the pilot study showed that words that seemed to be difficult for those speakers distracted them from producing the expected stress. Additionally, those participants tended to place more prominence on those supposedly difficult words to pronounce. In order to gain a greater understanding of the extent to which the complexity of words interfere with the production and perception of NS, this variable was controlled and investigated in the present study. For the listeners, understanding if they were acquainted with the vocabulary and with the pronunciation of the words would aid disentangle the reasons why they would misunderstand the speakers' intentions in the cases the NS was assigned to the expected portions of the speech. Completing these two tests took around 10 min for each pair.

While participants were completing the tests mentioned in the previous paragraph, the videos for Tasks 2, 3, and 4 were transferred to a laptop computer for later access during the interview with the listeners. After participants finished the tests, they were informed about the object of the present study and were asked if they had been taught about NS in their English classes.

After this informal talk, the speaker was asked to review all the tasks by reading each task and then underline the portions of the sentences (Tasks 2 and 3) and of the questions (Task 4) that they had to emphasize. This was asked in order to make sure that the speakers really understood the context in which the sentences and the questions were being uttered. After doing this last activity, the speaker was dismissed. Data collection for the speaker lasted at maximum 30 min.

The interview with the listener had the purpose of learning from the listener his or her impressions about the choices made in Tasks 2, 3, and 4. In order for easy future retrieval, this interview was audio-recorded. In the course of the interview, the sheet with the listeners' answers and the video-recordings were assessed. Data collection with the listener lasted at maximum 80 min. This data collection occurred from November 25 to December 15, 2015.

### **3.3.2 Summary**

This section described the instruments and procedures for data collection. The 14 pairs met the researcher and engaged in four pairwork

tasks: *Breaking the Ice* (Task 1), *I don't know the answer, do you?/I know the answer!* (Task 2), *You're wrong!* (Task3), and *I know what you're saying next!* (Task 4). Additionally, they took two tests: A *Word Familiarity Test* and a *Pronunciation Self-Evaluation Test*. In his/her sheets, the speakers underlined the words that should take the NS according to each discursive act and the listeners, in turn, were interviewed.

The data collected by means of the instruments and procedures described in Section 3.3 were analyzed as described in Section 3.4 in order to answer the research questions guiding the present study.

### 3.4 DATA ANALYSIS

Overall, the research questions and hypotheses in the present study, presented in Section 1.4, tackle issues of production (RQ 1-3) and interpretation (RQ 4-7). By confirming or disconfirming these hypotheses, it will be possible to have a minimum understanding of the way L1-BP speakers of English communicate with some members of the international community by means of NS assignment in order to elicit, correct, and contrast information.

For the analysis of the dataset, visual analysis and acoustic analysis were used to examine NS placement. The cue used for NS in the present study was pitch. Even though duration and loudness are also reliable cues reported in the literature, the design of the study did not favor the measurement of duration given that words had different length. However, an impressionist account of duration was done in order to shed light to some fuzzy results. Regarding loudness, only the perceived loudness when reported by the listeners were considered in the analysis. A more detailed analysis taking loudness into consideration is left for future studies. For the visual display of the F0 (fundamental frequency), which is the physical correlate of pitch, the audio recording files were open with *Praat* and the pitch curves were individually collected. For each utterance, the pitch curve was displayed in windows whereby F0 values were set near the highest and the lowest limits of the curve in order to keep the curve integrity, with no flattening, a procedure that is adopted in studies involving prosody (e.g., Carpes, 2014). To illustrate, consider Figures 9 and 10.

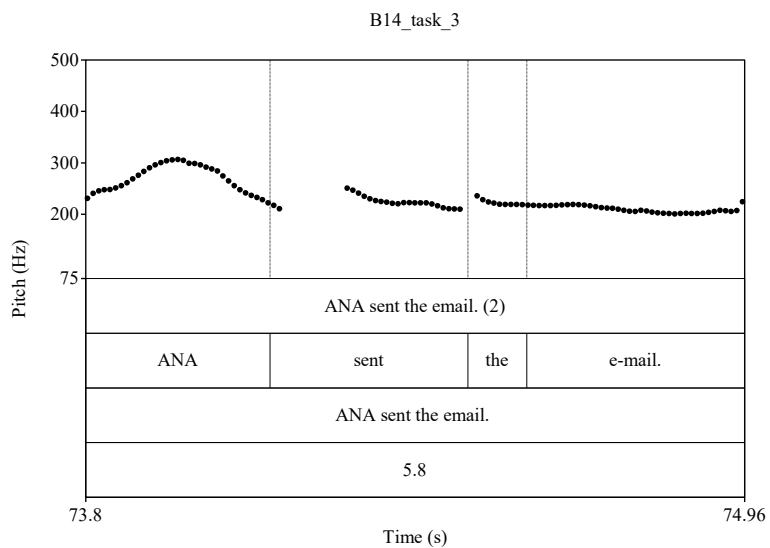


Figure 9: Random pitch settings.

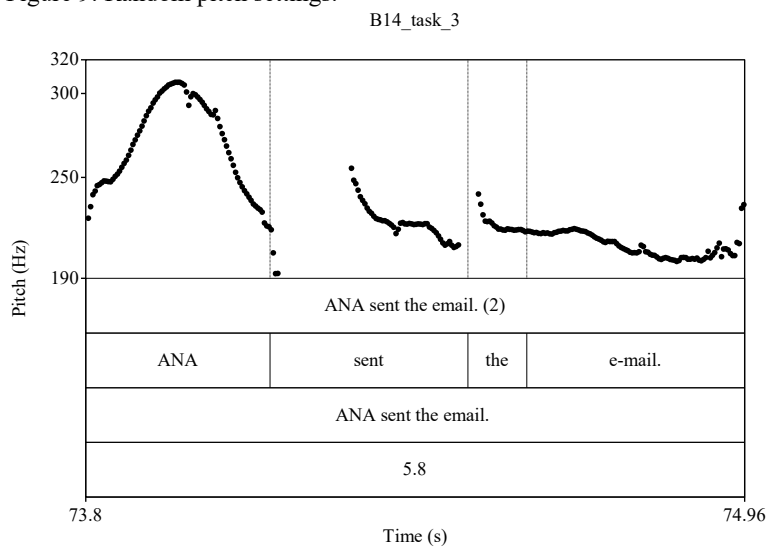


Figure 10: Pitch values set near the highest and the lowest limits of the curves.

Figures 9 and 10 show the production of *Ana sent the email* by Speaker B14. In Figure 9, the pitch settings range from 75 Hz to 500 Hz, the standards found in Praat. Because it is quite a wide range whose space is not occupied by the speakers' voice, the curves are flat. In Figure 10, the pitch settings range from 190 Hz to 320 Hz, a setting picked by the researcher as this range showed to respect the lowest and the highest limits of the curves for this specific production. Observe that this time the curve integrity is preserved. This analysis procedure was adopted for every production in the dataset<sup>52</sup>.

As argued by Krahmer and Swerts (2007), pitch change “needs to be above a perceptual threshold to be functionally relevant” (p. 121). According to t'Hart (1981), in order for variation in tones to be perceptually distinct, there is a need of a three-semitone change in pitch. While analyzing the speakers' productions in the pilot study, it was noticed that some pitch changes did not reach the three-semitone range. Moreover, sentences usually had more than one prosodic prominence. Based on the literature (e.g., Baptista, 2001), one of the hypothesis was that Brazilian speakers would place NS on its canonic position, that is, in the rightmost end of the sentence. Surprisingly, in the pilot study a great number of sentences had some prominence both at the expected location and at the leftmost edge of the sentences<sup>53</sup>. In light that measuring pitch in *semitones* rather than in *Hertz* enables a comparison between different pitch changes (Nootboom, 2007), the words in the sentences were segmented, labeled, the pitch changes were located, the vertical scales were transformed from Hertz into semitones, a tool offered by *Praat*, and the pitch change range was calculated for each pitch change location. The same procedures for analysis were used in

---

<sup>52</sup> It should be noted that each utterance was treated this way, irrespective of having been produced by the same speaker. That was so because range variation in productions were found for the same speaker. Therefore, the pitch setting was decided for each utterance individually based on the pitch tracking displayed during the analysis.

<sup>53</sup> Some of the utterances in the pilot study coincide with a phenomena described by Pierrehumbert (1980). When a speaker produces two thought groups, one made up by the subject of the sentence and the other made up by the action performed by the subject, it is likely it will be interpreted as a response to a doubled barreled question. To illustrate, consider the following data from the pilot study. *John went to Pennsylvania* produced into two thought groups “*JOHN % went to PENNSYLVANIA%*”. The barreled question being answered would be “Who went? And where to?” This may cause the listener some confusion because the listener was only questioning “Who went to Pennsylvania?”

the present study<sup>54</sup>. To illustrate the windows before and after transformation, consider Figures 10 and 11.

Every figure drawn for the purposes of the present study displays four tiers, as mentioned in Footnote 27 and rewritten here for easy retrieval.

- 1) Tier 1: Orthographic tier.
- 2) Tier 2: Word-boundary tier.
- 3) Tier 3: Thought-group-and-NS tier.
- 4) Tier 4: Pitch-change-range tier.

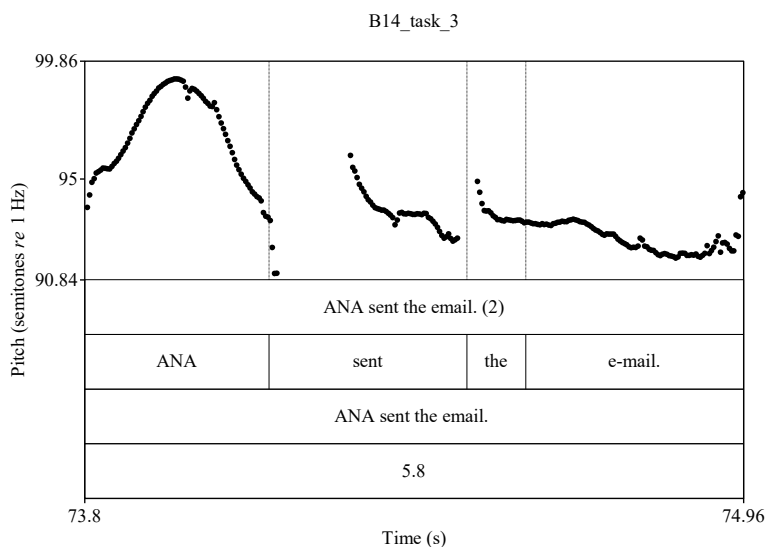


Figure 11: Praat window displaying vertical scales transformed from Hertz to semitones.

In Figure 11, Tier 1 displays the utterance to be produced. Tier 2 displays the limits of each word. Tier 3 shows the utterance with the thought groups limits and the constituents with distinctive pitch change. Finally, Tier 4 includes the total number of semitones in each pitch

<sup>54</sup> During data analysis of the production data, it was noted that Praat could not track the pitch properly in some cases and the research had to rely on her perception in order to make a decision of which track to follow. For examples, refer to Appendix R.

change. A comparison of Figures 10 and 11 shows no differences in terms of the curves displayed. The only difference lies on the possibility for the researcher to measure the pitch change in semitones. Note that *Ana* is the location where more pitch change is found, i.e., a change of 5.8 semitones (st), while in the remainder of the production the change was not great. This signals that the speaker chose to nuclear-stress *Ana* over *sent the email*, meaning that the important information here is the agent rather than the action. For utterances produced with more than two pitch changes, only pitch differences of more than 1.5 st were considered to create reliable differences for the perception of NS (Rietveld & Gussenhoven, 1985 as cited in Nootboom, 1997).

The thought groups within each utterance were identified according to the number of pauses made by each speaker. In the literature, no consensus has been reached regarding how long a period of silence has to be in order to be considered a pause, which has ranged from 100 to 400 milliseconds (ms) of threshold (Lege, 2012). On the one hand, the course of pitch can be interrupted during the production of voiceless consonants such as [p], [t], and [k], due to articulatory reasons, but this interruption is not heard in fluent speech if the interruption is shorter than 200 ms (Nootboom, 2007; Warren, 2013). This lack of perception of the articulatory silence in the course of pitch is granted to the fact that only “silent gaps longer than this [200 ms] effectively prohibit perceptual integration of preceding and following speech sounds” (Nootboom, p. 5). Therefore, researchers agree to set a duration of 200 ms in order for silences to be considered as pauses (Warren, 2013). On the other hand, pauses are noticed even when there is no silence. Other signals such as duration of syllables and pitch change are cues to identify them (Pierrehumbert, 1980). Based on this, silences longer than 200 ms were considered as pauses and silences shorter than that associated with rising intonation and/or lengthening of syllables were also considered as indicators of pauses in speech. Those pauses were signaled by means of a hyphen ( - ) in the annotation tiers in each figure generated during data analysis. Gaps that were not considered as pauses did not include a hyphen. Thought groups were considered as having a NS only if differences in the lower pitch (valley) and the higher pitch (peak) (and vice-versa) were equal to or higher than 3 st<sup>55</sup>. Figure 12a displays an example of productions with pauses,

---

<sup>55</sup> The measurement of semitones was carried out by checking the lowest pitch near the context of pitch change, irrespective of being within the portion in focus (Carpes, 2014).



yielding more than one thought group and thus more than one NS in a given production.

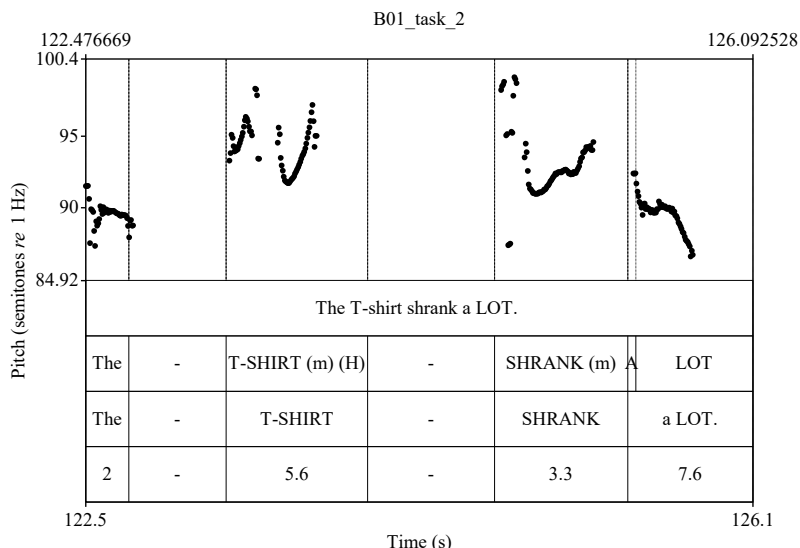


Figure 12a: Production of The T-shirt shrank A LOT by speaker B-01, an illustration of an utterance with more than one thought group caused by pauses.

Figure 12a displays (Tier 1) the sentence produced, (Tier 2) the words and pauses limits, (Tier 3) the NS and thought groups, and (Tier 4) the pitch change range in semitones. Observe that this production was broken into four thought groups, three separated by pauses: (1) *The*, (2) *T-shirt*, and (3) *shrank*, and one separated by rising intonation (4) *a lot*. The durations of the pauses were 509 and 646 ms, respectively. The word *The* did not show a distinctive pitch change (2 st), while the other thought groups did (a change of 5.6, 3.3, and 7.6 st, respectively). Therefore, three distinct nuclear stresses were found in this production. In relation to the pauses, note that before the pause limits after *shirt*, *shrank*, and *lot* there are also blank gaps (with no pitch tracking). However, if the spectrogram was also visible, the waves would be seen in the blank spaces in Figure 12a, as seen in the replication of the same production in Figure 12b. Because this study was concerned with the pitch tracking, only the fundamental frequency tracking was maintained in the figures illustrating the productions.

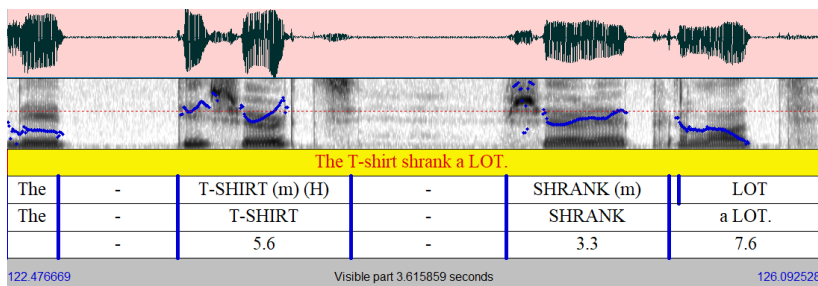


Figure 12b: Production of *The T-shirt shrank A LOT* by speaker B-01, an illustration of wave portions that were not pitch tracked.

In order for a production to be considered produced as expected in terms of NS assignment, it had to meet two requirements: (1) having only one thought group and thus only one NS, and (2), having the NS placed at the expected portion of the utterance as established by each discursive context. The productions that did not meet these requirements were considered as unexpected productions. The information obtained by means of these procedures for analysis were coded and transferred to an SPSS spreadsheet so that descriptive statistics and statistical tests could be run in order to answer the research questions related to production (RQ1, RQ2, and RQ3).

In order to look into the way these productions were interpreted with relation to the intent of the speakers (RQ 4, RQ 5, RQ 6, and RQ7), the perception of NS was verified and classified as expected and unexpected. A perception of NS was considered as expected if the listener managed to perceive the NS at the expected location. The expected perception of NS equals successful communication as it led to the expected interpretation of the speaker's intent.

The unexpected perceptions were classified into two categories: (1) sites of reduced interpretability and (2) real communication breakdown. Sites of reduced interpretability are defined here as any pause group containing a section of speech where a listener was unable to, or had difficulty in identifying the speakers' intent (Zielinski, 2008), making communication between speaker and listener hard, but successful after some meaning negotiation. These sites were signaled by means of need for repetition and the consequent expected perception of NS in the repeated production, which did not lead to total failure in communication. Therefore, reduced interpretability equals to noise in communication without full communication breakdown.

An unexpected NS perception was considered to depict a real communication breakdown in two different cases: (1) when it led to an unexpected interpretation of the speaker's intent based only on the first production (no repetition required), and (2) when it led to an unexpected interpretation of the speaker's intent based on the speaker's second or third production (when repetition was required). Therefore, real communication breakdown depicts two different sides of unsuccessful communication. On one side, when no repetition occurs, that is, no negotiation of meaning happens, neither the speaker nor the listener is aware that the speaker's intent was not interpreted as intended by the speaker. On the other side, when there is repetition, both speaker and listener are aware of the fragility of that communication act and that failure might have happened.

In the present dataset, pieces of evidence for unexpected perception, sites of reduced interpretability and real communication breakdowns were collected from (a) the listeners' behavior while listening; (b) the listeners' answers to the questions in the task (matching or not the expected interpretation according to the discursive context), (c) the need for repetition (which signaled that the listeners were unsure about the speakers' intent), and (d) the listeners' comments in the follow-up interview. Additionally, their answers to the *Word Familiarity Test* and the *Pronunciation Self-Evaluation Test* were considered in order to help explain unexpected behaviors. Finally, the videos were also watched in order to check if NS production or perception coincided with extra visual clues, such as eyebrow and head movements. These pieces of information were coded and transferred to an SPSS spreadsheet so that descriptive statistics and statistical tests could be run.

Regarding the productions considered to answer the research questions related to the interpretation of the speakers' intent (RQ from 4 to 7), a decision was made of dismissing 14 of the 378 utterances produced for a couple of reasons: (1) the speaker said more than she was expected, and (2) the speaker did not understand the context and thus misunderstood which portion of the utterance the NS should fall. Table 12a displays the dismissed utterances.

Table 12a  
*Dismissed cases*

Speaker	Utterances	Production		Interpretation		
		Exp	Unexp.	Exp	Unexp.	
B01	Do MEN lie a lot?		1		1	
	Is ZENNY afraid of heights?		1		1	
	Does the meeting START at ten?		1		1	
	Does this sequence REPRESENT the hierarchy...?		1	1		
	Does the party start at TEN?		1		1	
	Did Ana twirl her HAIR?		1	1		
	Is the atmosphere around Zenny GOOD?		1	1		
B02	Does the meeting START at ten?		1		1	
B05	ZENNY sent you the pearls.		1		1	
	Zenny loves DOGS. <sup>a</sup>		1		1	
B11	Does this sequence REPRESENT the hierarchy?		1		1	
B13	26 Did Ana twirl her HAIR?		1	1		
B14	Does this sequence REPRESENT the hierarchy?		1		1	
B15	Does this sequence REPRESENT the hierarchy?		1		1	
Total		14	0	14	4	10
			0%	100%	29%	71%

<sup>a</sup> The speaker said more than she should.

As illustrated in Table 12a, 14 productions were dismissed: one because the speaker said more than she should (B05) and 13 because the speaker did not understand the context properly in order to place the NS at the expected portion of the utterance, according to their highlights in their sheets after performing the tasks. All utterances had unexpected productions. However, 28.6% (N = 4) of these productions were interpreted as intended when heard for the first time. In order to see illustrations of the dismissed cases, please refer to Appendix S.

### 3.5.1 Statistics

Descriptive statistics were run with the help of *SPSS Statistics 17.0*, for Windows. Table 12b summarizes the dependent and independent variables, and their levels for each research question in the present study.

Table 12b

*Summary of the variables and their levels according to each research question*

Research Questions	Dependent variable	DV levels	Independent variables	IV levels
RQ1	NS placement	1) Expected 2) Unexpected	Position	1) initial 2) medial 3) final
RQ2	NS placement	1) Expected 2) Unexpected	Type of information	1) being elicited 2) being corrected 3) being contrasted
RQ3	NS placement	1) Expected 2) Unexpected	Complexity of words	1) non-challenging 2) short-challenging 3) long-challenging
RQ4	NS perception	1) Expected 2) Unexpected	NS production	1) Expected 2) Unexpected
RQ5	NS perception	1) Expected 2) Unexpected	Position	1) initial 2) medial 3) final
RQ6	NS perception	1) Expected 2) Unexpected	Type of information	1) being elicited 2) being corrected 3) being contrasted
RQ7	NS perception	1) Expected 2) Unexpected	Complexity of words	1) non-challenging 2) short-challenging 3) long-challenging

As can be seen from Table 12b, there are two dependent variables: *NS production* and *NS perception*. *NS production* is investigated in research questions 1, 2, and 3. *NS perception* is investigated in research questions 4, 5, 6, and 7. Both dependent variables are categorical and have two levels: *expected* and *unexpected*.

The independent variables are *expected NS position*, *type of information*, and *complexity of words* for both dependent variables and they regard research questions 1 and 5, 2 and 6, 3 and 7, respectively. They all have three levels. The independent variable *NS production* regards research question 4 and it has only two levels. All the independent variables are categorical.

The chi-square test of independence was chosen for testing the statistical significance for three main reasons: (1) all variables, dependent and independent, are categorical, (2) it is not possible to calculate averages of the variables, that is, it is only possible to count how many are in each category, and (3) each variable has two or more levels within them (Larson-Hall, 2010).

Except for Research Question 4, the conventional way of running chi-square could not be used due to the fact that the independent variables had three levels. Therefore, in order to test the statistical significance of the differences in frequencies found between the three levels of the independent variables, an advanced way of Chi-Square for testing independence was used (Larson-Hall, 2010). This way consists of transforming a 2 x 3 contingency table into two 2 x 2 contingency tables. This partitioning is demonstrated in detail in Section 4.1.1.

According to Larson-Hall (2010), the easiest way to perform the test without having to rearrange the data in SPSS is to do the testing using an online calculator. The present study followed Larson-Hall's suggestion<sup>56</sup>. The procedures for using the calculator are explained in detail in Section 4.1.1.

Even though Larson-Hall's (2010) method for Chi-Square was employed, I opted for running multiple comparisons and pairing the results found for the two partitioned 2 x 2 contingency tables with the ones found in the multiple comparisons. The probability level of statistical significance was set at .05 for the partitioned tables and at .01 for the multiple comparisons (Bonferroni Correction) using the same online calculator.

---

<sup>56</sup> The online calculator available at <http://www.graphpad.com/quickcalcs/Contingency1.cfm> was used.

### 3.5.2 Summary

This section explained the procedures for data analysis. Each production was displayed at a Praat window, the words and pauses were segmented, and the pitch changes were mapped so that the NS assignment was identified. The listeners' behavior, answers, tests and interview were assessed in order to understand the way these listeners interpreted the speakers' intent. Chi-square was the statistical test used to verify if the results reached statistical significance.

### 3.6 CONCLUSION OF THE CHAPTER

First, the present chapter provided the steps for recruiting and selecting the participants. Their English and Portuguese classes were visited and they were invited to participate. Only the ones willing to participate, whose schedule availability matched the laboratory's working hours and whose level of oral proficiency were intermediate or higher (as described in Sections 3.1.1, 3.1.2, and 3.1.3) participated in the study.

Second, the chapter introduced the participants in the study, 14 L1-BP speakers of English and 14 foreigners, by providing some background data and some information on their experience with English learning and English use (Sections 3.2.1 and 3.2.2). Speakers and listeners differed mainly with regards to (1) time and place when/where they started/learned their English and (2) their English usage rate/level of proficiency.

Third, the chapter covered the procedures and instruments for data collection: four audio-and-video-recorded-pairwork activities, a written word familiarity test, a written pronunciation self-evaluation test, and an audio recorded interview, as described in Section 3.3.

Fourth, the objective of the study, the seven research questions guiding it and the seven hypotheses were introduced. Three hypotheses made predictions about the production of NS by the speakers while the other four made predictions about the way their intent would be interpreted (Section 1.4). In general terms, it was predicted that unexpected assignment of NS would lead to unexpected interpretation of speakers' intent.

Finally, the procedures for data analysis were introduced and explained in Section 3.4. Some attention was given to the pitch settings in Praat due to its importance in guaranteeing an accurate display of pitch contour for the visual identification of NS. Additionally, instead of using *Hertz*, *semitones* were used in order to measure the changes in pitch. The criteria for setting the limits of thought groups were also listed: pauses higher than 200 ms, syllable lengthening and/or rising intonation. Section 3.4 also explained the criteria used to consider a production as expected or unexpected and an interpretation as expected or unexpected. The next chapter provides detailed information about the results and discusses the findings of the present study.



## 4 RESULTS AND DISCUSSION

This study is interested in understanding how English-speaking members of the international community interpret Brazilians' intentions according to nuclear stress (NS) placement to emphasize important information. Thus, as discussed in the previous chapter, the data collected were a result of the interaction in English of Brazilians with members of the international community. The present chapter is dedicated to presenting the results of data analysis and discussing the findings. In an attempt to make the chapter easier for the reader to follow, this chapter writing is oriented according to the research questions listed in Chapter III.

Section 4.1 introduces and discusses the production results (RQ1, RQ2, and RQ3). It starts by reviewing the conditions that need to be met in order for a NS placement to be considered or not an expected production and by presenting the general production results when utterances were produced only once and when they were repeated. Section 4.1.1 addresses RQ1, by presenting the production results according to NS location in the utterance and the statistical test (Chi-square test for group independence) to confirm or not H1. Additionally, it contains a detailed account of the productions and some illustration by means of Praat window figures. Section 4.1.2 addresses RQ2, by presenting the production results according to NS in utterances according to the types of information and the statistical test (Chi-square) to confirm or not H2. Section 4.1.3 addresses RQ3. It starts by reviewing the utterances produced in the present study. Next, the section presents the results for the Pronunciation Self-Evaluation Test taken by the speakers. After that, it reports on the production results according to the complexity of words in the utterances and the statistical test results to confirm or not H3, and provides some illustrations. Section 4.1.4 summarizes Section 4.1.

Section 4.2 introduces and discusses the perception results (RQ4, RQ5, RQ6, and RQ7). It starts by reviewing the roles played by the listeners during the tasks and the two levels of perception: expected and unexpected. It also (1) reviews the concepts of successful interpretations (related to expected NS perception), sites of reduced interpretability and real communication breakdowns (related to unexpected NS perception), and (2) presents the general perception results. Section 4.2.1 addresses RQ4, by reporting on the results for perception of NS according to the

expected or unexpected NS placements. It also includes the explanations reported by the listeners for interpreting or not as expected the speakers' intents (with illustrations provided). Section 4.2.2 addresses RQ5 and is divided into two other sections. Section 4.2.2.1 answers RQ5 by reporting on the perception results based on the expected NS location and the results for the statistical test. Section 4.2.2.2 presents the analysis of data regarding the perception of speakers' actual NS placement and the resulting interpretation of speakers' intent, by drawing detailed comparisons between the production cases found in Section 4.1.1 and their perception. Section 4.2.3 addresses RQ6, by reporting the perception results for NS in utterances conveying the three types of information investigated and the statistical test results in order to confirm or not H6. Section 4.2.4 addresses RQ7. It is divided into two sections. Section 4.2.4.1 answers RQ7 by reporting on the perception results based on the presence of complex words in the utterances and the results for the statistical test. Section 4.2.4.2 presents the analysis of data regarding the perception of speakers' actual NS placement and the resulting interpretation of speakers' intent, by drawing comparisons between the production and perception according to the complexity of words. Additionally, it addresses the effects of the combination of two or more of the pronunciation/prosody production aspects listed in Appendix T on perception. Finally, Section 4.2.5 summarizes Section 4.2. In the beginning of Sections 4.1.1 to 4.2.4, the research questions and hypotheses corresponding to each section are repeated for easy retrieval.

Lastly, Section 4.3 concludes the chapter, by revising the main findings for NS placements by L1-BP speakers of English and for the NS perception and speakers' intent interpretation by IL-English listeners.

#### 4.1 BRAZILIANS' PLACEMENT OF NS IN ENGLISH

This section presents and discusses the production results for NS placement in English by Brazilians. Fourteen Brazilians were paired with 14 members of the international community and produced 378 utterances and 123 repetitions ( $N = 501$ ) that were analyzed according to the placement of NS to convey specific information in given discursive contexts. The 501 utterances were acoustically analyzed as explained in Section 3.4. Two utterances were repeated twice, but only the second

repetition was computed for analysis as we suppose the listener would base his or her final answer on the second rather than on the first repetition of the utterance. This being so, 121 repetitions were taken into consideration in the results, summing up 499 productions.

A production was considered as expected when it had only one thought group and thus only one NS in the utterance placed at the expected portion of the utterance. Productions were considered unexpected in the cases described below:

- 1) The production had one thought group and one NS placed in an unexpected portion of the utterance;
- 2) The production had two or more thought groups, with two or more NS positions, including or not the expected one.
- 3) The production had no distinct pitch change and thus no NS.

Table 13 shows the general production results for the 378 utterances when produced for the first time and the 121 repetitions.

Table 13  
*NS placement by Brazilians in English – general results*

Status	First-time production		Repetition		Total	
	N	%	N	%	N	%
Unexpected	315	83.3	101	83.5	416	83.4
Expected	63	16.6	20	16.5	83	16.6
Total	378	100.0	121	100.0	499	100.0

As can be seen, 83.3% of the utterances (N = 315) had an unexpected production when said for the first time in terms of NS placement. When repeated, 83.5% of the utterances (N = 101) were still not produced as expected, signaling that the repetition did not help in most cases the expected placement of NS.

The sections that follow discuss the results according to NS location (Section 4.1.1), type of information (Section 4.1.2), and level of difficulty of words in terms of pronunciation (Section 4.1.3), which are related to research questions 1, 2, and 3.

#### 4.1.1 RQ1 – Production results for location of NS

RQ1: How is NS placed by L1-BP speakers of English, according to the expected position of NS (initial, medial, final) in utterances?

H1: Based on the findings in the pilot study (Passarella-Reis & Silveira, 2016), Hypothesis 1 predicts that the position of NS in the utterance will not interfere with either the expected or unexpected production of NS.

NS position in the utterance was an independent variable in the present study investigated in RQ1. Although this was not a variable formally investigated in the pilot study, data analysis of the pilot study showed that NS position in the utterance was not a variable that seemed to affect the expected placement of NS. Therefore, although the canonical position for NS is the final portion of the utterance (Lieberman, 1975), Hypothesis 1 predicted that the fact that one has to place NS in a place other than the final portion in an utterance, as set by the discursive contexts, would not interfere with the expected placement of NS. Tables 14 and 15 show the results according to expected NS position: initial, medial, or final for first-time and repeated productions, respectively.

Table 14

*Production results according to NS location – first-time production*

Status	Initial		Medial		Final		Total	
	N	%	N	%	N	%	N	%
Unexpected	98	77.8	103	81.7	114	90.5	315	83.4
Expected	28	22.2	23	18.3	12	9.5	63	16.7
Total	126	100	126	100	126	100	378	100.0

Table 15

*Production results according to NS location - repetition*

Status	Initial		Medial		Final		Total	
	N	%	N	%	N	%	N	%
Unexpected	23	69.7	40	88.9	37	86	100	82.6
Expected	10	30.3	5	11.1	6	14	21	17.4
Total	33	100	45	100	43	100	121	100.0

When utterances were produced for the first time (Table 14), results show that NS was successfully placed more frequently in initial position (22.2%), followed by medial position (18.3%) and then by final

position (9.5%). When the utterances were repeated (Table 15), it was still placed as expected more often in initial position (30.3%), but this time followed by final position (14%) and then by medial position (11.1%).

One should expect that the allocation of NS on the final portion of the utterance should be easier because it is the canonical location for NS in neutral contexts (Liberman, 1975). However, it showed to be slightly more difficult than it is when in the medial and the initial portion of the utterance. The frequencies were tested in order to check if the differences were statistically significant. A Chi-Square independence test would be used if there were only two levels (2 x 2) in the variables. However, position has three different levels: initial, medial, and final. Therefore, a regular Chi-Square would not suffice, as it deals with only two levels. Larson-Hall (2010) suggests that data be rearranged by partitioning the contingency table in order to have two 2 x 2 tables instead of an only 2 x 3 table so the Chi-square test can be run. Tables 16 and 17 show the rearranging of data and the partitioning, respectively.

Table 16

*Production results rearranged for statistical testing (2 x 3 Contingency Table) – IV Position*

Position	Expected		Unexpected	
		N		N
Initial		28		98
Medial		23		103
Final		12		114
Total		63		315

Table 17

*Production results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Position*

Position	Expected	Unexpected	Position	Expected	Unexpected
	N	N		N	N
Initial	28	98	Initial + Medial	51	201
Medial	23	103		Final	12
Total	51	201	Total	63	315

With the partitioning, the first two rows (initial and medial) and the two columns (expected and unexpected) of Table 16 generated the first 2 x 2 contingency table in Table 17 (displayed on the leftmost half). The next partition generates the second 2 x 2 contingency table in Table 17 (displayed on the rightmost half) by combining the first two rows (initial + medial) and by comparing them to the third row (final) of Table 16. Thus, two partition tables are created to test with the group-independence chi-square test. According to Larson-Hall (2010), the easiest way to the testing without having to rearrange the data in SPSS is to do the testing using a quick online calculator. Larson-Hall suggests the one available at <http://www.graphpad.com/quickcalcs/Contingency1.cfm>. On this webpage, the data in the initial and medial rows are punched in and the option “Chi-square without Yates’ correction” has to be chosen, as advised by Larson-Hall, even though the website tells us to use Fisher’s exact test. Additionally, the website recommends that the P value be calculated using two-tailed P values and this is checked by default. I just left it the way it was, as no mention was made to it by Larson-Hall. Next, I pushed the “calculate” bottom and the website opened another webpage with the test results, which showed no statistical significance between these variables (NS placement in initial and medial positions):  $X^2(1, N = 378) = 0.615, p = .4331$ . This means that L1-BP assignment of NS is not affected by the position of NS when in initial or medial position, being equally challenging in both.

The next step was to repeat the same procedures for initial + medial and final rows of Table 17. The test showed that the relation between these variables is statistically significant:  $X^2(1, N = 378) = 6.943, p = .0084$ . This means that placing NS in final position showed to be more difficult for L1-BP speakers of English than it was for them to place NS in initial and medial position and this difference was statistically significant.

Because I thought that having multiple comparisons would make the relations a bit more clear, instead of testing the data in the partition contingency tables, I opted for testing each variable against each other. For doing so, as there are three levels and multiple comparisons, the  $p$  value was recalculated with Bonferroni correction. Therefore, a  $p$  equals .01 or smaller would show a statistically significant difference. With these new calculations, the difference was significant only between final and initial position:  $X^2(2, N = 378) = 6.686, p = .0097$ . It approached

significance between final and medial positions:  $X^2(2, N = 378) = 3.318, p = .0685$ , while it was not statistically significant for initial and medial positions:  $X^2(2, N = 378) = 0.281, p = .5959$ .

Hypothesis 1 predicted that NS position in the utterance would not play a role in the production of NS by L1-BP speakers of English, as they would show difficulties to place NS irrespective of utterance position, but it was not confirmed. Most utterances had unexpected NS placements, but utterances with NS expected to be placed in final position were the most challenging. It is a surprising finding given that the final is the canonical NS position in wide-focus contexts (Liberman, 1975) and that previous studies have found that IL-English speakers tend to place NS at the rightmost end of utterance even in narrow focus contexts (e.g., Tiffen, 1974). However, if one refers back to the literature, one finds some indication that L1-BP speakers of English tend to place NS at the rightmost end of English wh-questions and some pitch movement is also found in initial position (Manosso, 2013), making arise two potential places for NS, which does not match an expected use of NS in English. Unfortunately, wh-questions were not a structure used in the present study, and a direct comparison cannot be drawn. Research has also found that, in BP polar questions (Moraes et al., 2015) and utterances (Carpes, 2014), when NS falls on the rightmost end of the utterances, L1-BP speakers tend to make pitch changes in other positions (Moraes et al., 2015), mainly in initial position (Carpes, 2014). It may be the case that these speakers transferred this NS use to English, which might explain why NS on final position had the most unexpected productions, as pitch movement was also found somewhere else in the utterances, which is not a trait of English. In BP utterances, research has also found that L1-BP speakers use some destressing when NS is on initial position (Carpes, 2014). It may be the case that these speakers transferred this destressing in some productions, which might explain why NS in initial position was placed as expected a little more often than it was in final position.

In order to understand the way these utterances were produced when the expected production was not reached, I examined the unexpected production in relation to the location of NS. Table 18 shows the production details for the first time these utterances were produced and Table 19 displays the ones for the times they were repeated. These tables are organized as follows: the rows displays the NS production details and the columns the frequency for each level of the independent

variable (position – initial, medial, final). The row labeled “*Expected*” displays the frequencies for the expected NS placements, while the remainder of the rows displays the frequencies for unexpected placements as follows: the row labeled “*As initial*” includes the frequencies when NS was misplaced to initial position; in the row labeled “*As medial*” there are the frequencies when NS was misplaced to medial position; in the row labeled “*As final*” there are the frequencies when NS was misplaced to final position; the row labeled “*Two+ locations, + expected*” includes the frequencies for NS placed at two or more locations with a greater pitch change range on the expected location; the row labeled “*Two+ locations – no ≠*” displays the frequencies for NS placed at two or more locations with no distinct differences between pitch change ranges; the row labeled “*Two+ locations, + unexpected*” includes the frequencies for NS placed at two or more locations with a greater pitch change on an unexpected location; and the row labeled “*No NS*” displays the frequencies for productions in which pitch changes did not reach the 3-st necessary change.

Table 18  
*Production details according to NS location – first-time production*

Details	Initial		Medial		Final		Total	
	N	%	N	%	N	%	N	%
Expected	28	22.2	23	18.3	12	9.5	63	16.6
As initial	0	0	2	1.6	5	4.0	7	1.9
As medial	2	1.6	0	0	3	2.4	5	1.3
As final	1	.8	3	2.4	0	0	4	1.1
Two+ locations, + expected	37	29.4	18	14.3	34	27.0	89	23.5
Two+ locations, no ≠	32	25.4	43	34.1	44	34.9	119	31.5
Two+ locations, + unexpected	25	19.8	31	24.6	24	19.0	80	21.2
No NS	1	.8	6	4.8	4	3.2	11	2.9
Total	126	100	126	100	126	100	378	100.0



Table 19

*Production details according to NS location – repetition*

Details	Initial		Medial		Final		Total	
	N	%	N	%	N	%	N	%
Expected	10	30.3	5	11.1	6	14.0	21	17.4
As initial	0	0	0	0	2	4.7	2	1.7
As medial	0	0	0	0	1	2.3	1	.8
As final	1	3.0	2	4.4	0	0	3	2.5
Two+ locations, + expected	8	24.2	7	15.6	9	20.9	24	19.8
Two+ locations, no ≠	10	30.3	16	35.6	13	30.2	39	32.2
Two+ locations, + unexpected	4	12.1	14	31.1	12	27.9	30	24.8
No NS	0	0	1	2.2	0	0	1	.8
Total	33	100	45	100	43	100	121	100

As can be seen from the tables, not many utterances were produced as expected both when uttered by the first time (16.9%, N = 63) and when repeated (17.4%, N = 21). Figure 13 illustrates an expected production.

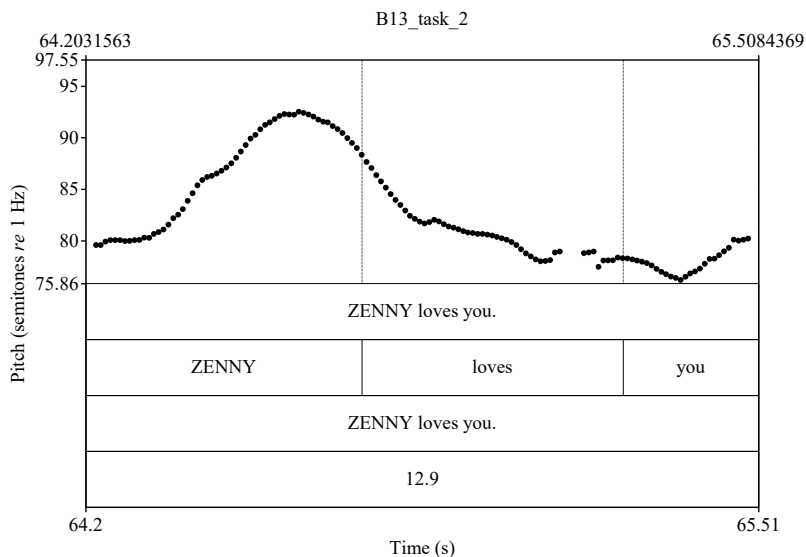


Figure 13: *ZENNY loves you* (Task 2, Item 04) by B13 – An example of an expected NS placement.

Figure 13 has four tiers, as all figures generated with Praat in the present study as explained in Chapters 2 (Section 2.1.5.2.2) and 3 (Section 3.4). Recall that for a pitch change to be considered a NS it was necessary that the change in pitch was greater than three semitones. The context set for this utterance was a conversation between two coworkers. One coworker had heard that someone loved her and asks who loves her. The speaker, the other coworker in the conversation, knows who loves her and needs to say the utterance with the emphasis on *Zenny*. In B13's production of *ZENNY loves you* (Figure 13), note that the pitch change has 12.9 st and is located at the word *Zenny*, the expected place for NS, meaning that the speaker wanted to inform who the person that loves the listener is, as set by the discursive context.

There were cases in which an only thought group was produced, but the NS was not placed at the expected location (initial, medial or final). It happened 4.3% (N = 16) of the times when produced for the first time and 5% (N = 6) when repeated. Figure 14 illustrates this type of production.

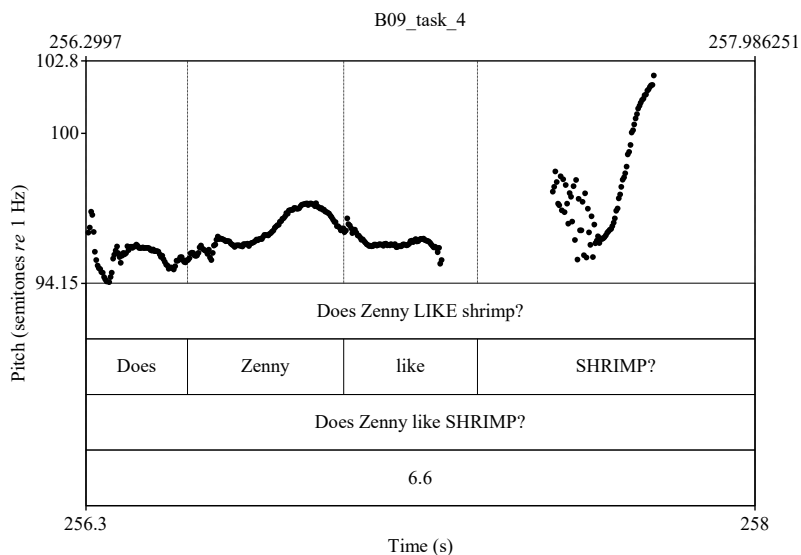


Figure 14: *Does Zenny LIKE shrimp?* (Task 4, Item 09) by B09 – An example of a NS placed on a position other than the expected.

In Task 4, the speakers had to ask only the first half of a disjunctive question. In the case represented in Figure 6, the contrast should be placed on *like* to signal that the speaker was in doubt if Zenny liked or disliked shrimp so that they could include it or not in the menu for Zenny's surprise party. However, in B09's delivery of this production, the NS fell on *shrimp* rather than on *like*, implying that she was in doubt about the kind of food that Zenny liked: shrimp or salmon. Note that, although there is some pitch change on *Zenny*, it is lower than three semitones in range (95 to 97.3 st) and thus it was not considered a NS. It might be that B09 misplaced NS in this specific case because she was more concerned with making the utterance sound like a question and pronouncing *shrimp* intelligibly, which by the way she did, than to placing the NS on the expected position.

Many utterances were produced with two or more thought groups and thus had two or more NS locations. These were classified into three categories: (1) NS with greater pitch change at the expect portion of the utterance; (2) NS with no distinction in pitch changes; and (3) NS with greater pitch change at an unexpected portion of the utterance. In the first-time production, 23.5% (N = 89) of the utterances were produced

as in case (1), with a greater pitch change on the expected portion of the utterance, and when repeated, 19.8% (N = 24) were produced under the same fashion. Figure 15 displays one example of these productions.

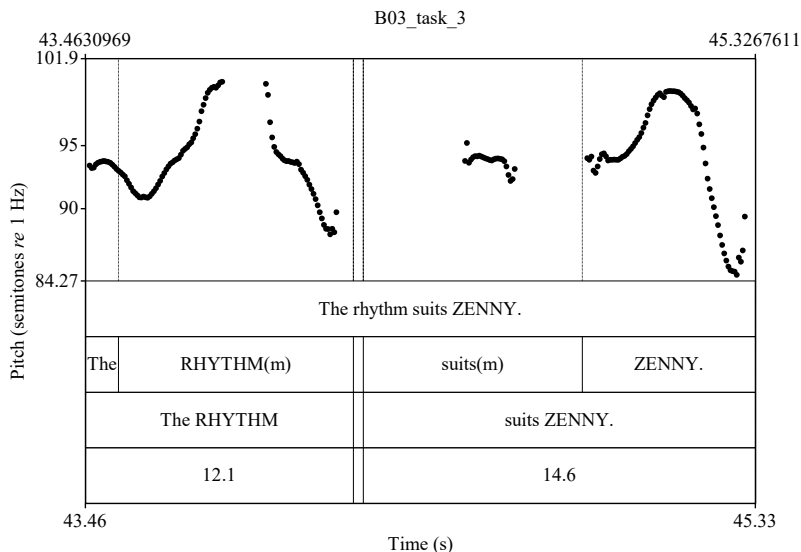


Figure 15: *No, the rhythm suits ZENNY* (Task 3, Item 03) by B03 – An example of NS placed in more than one location with a greater pitch change range on the expected portion of the utterance.

In Task 3, the speakers had to place NS to correct information. In this specific example displayed in Figure 15, B03 was expected to nuclear-stress *Zenny* to correct the information that the rhythm suited Ana rather than Zenny, when choosing the best rhythm to be played at the surprise party for Zenny. She is able to place the NS on *Zenny* with a 14.6-semitone-pitch change. Nevertheless, she also produced a pitch change in the beginning of the utterance, probably due to her difficulty to pronounce *rhythm*, released as [*ˈrɪθmi*]. Therefore, her production had two thought groups: (1) *The RHYTHM*; and (2) *suits ZENNY*. This specific production yielded the expected interpretation (for information on the expected interpretations, refer to Section 4.2), probably due to the combination of NS, eyebrow movement, and a slight nod of the head<sup>57</sup>,

<sup>57</sup> As observed in the videos recorded.

extra cues that have proved to aid the perception of NS (Keating et al, 2003).

Productions with two or more NSs with no differences in pitch changes, case (2), abounded even more in the dataset. In the first-time production, they represent 31.5% (N = 119) of the utterances and when repeated, 32.2% (N = 39). A change in pitch was considered different from one another only if they had at least a 1.5-semitone-pitch-change difference (following Rietveld & Gussenhoven, 1985 as cited in Nootboom, 1997). Therefore, productions in case (2) did not reach that pitch change difference. Figure 16 illustrates this sort of production.

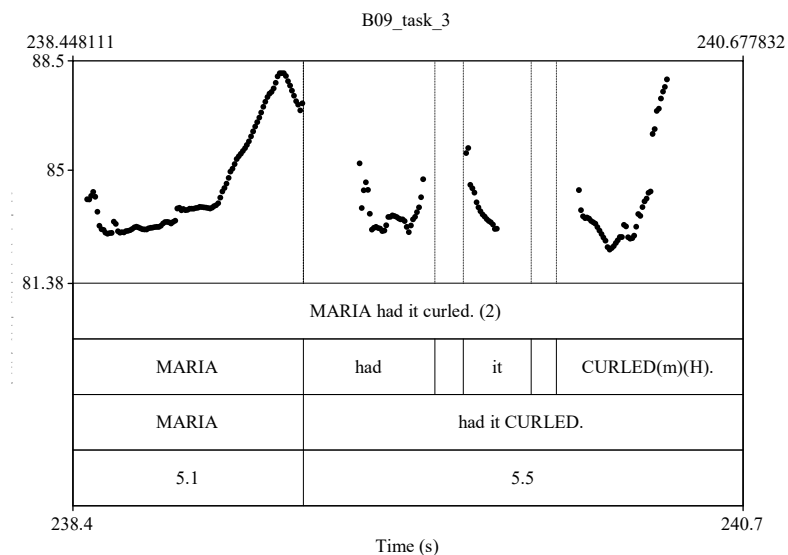


Figure 16: *No, MARIA had it curled* (Task 3, Item 09) by B09 – An example of NS placed in more than one location with no distinction in pitch change ranges.

The context for the production in Figure 16 was an informal conversation about who had the hair curled at the beauty salon. B08 had in the context sentence the information that *Ana had her hair curled* and was expected to correct that *Maria* rather than *Ana* had the hair curled, by placing NS on *Maria*. B08 successfully placed NS on *Maria*, with a pitch change of 5.1 semitones. However, there is another pitch change in *curled*, pronounced [ˈkɜrlɪd] with an unexpected rising intonation of 5.5

semitones. Therefore, this production had two thought groups and two NSs: (1) *MARIA* and (2) *had it CURLED*. The NS placement in *curled* was probably a result of B09's uncertainty about how to pronounce the word, even though she rated her pronunciation for this word as being intelligible enough. Even with the word *Maria* having a similar pitch change to that of *curled* and being produced in tandem with a slight eyebrow movement, it was not perceived as holding the NS and thus this production had an unexpected interpretation, as F09 perceived the NS on *curled*. F09 reported that the short preceding pause<sup>58</sup> (83ms) and the way *curled* was pronounced caught his attention<sup>59</sup>.

Finally, case (3), in which two or more NSs were produced with a greater pitch change at an unexpected portion of the utterance, was also abounding: 21.2% (N = 80) of the utterances in the first-time productions and 24.8% (N = 30) in the repetitions. Figure 17 provides an illustration of this case.

---

<sup>58</sup> It should be noted that the perceived pause before *curled* could have been only an articulatory gap due to the presence of the voiceless stops [k] and [t]. It is interesting to note, however, that this silence gap was shorter than 200 ms (the measurement made considering not only the F0 tracking but also the waves displayed in the spectrogram, as shown in Figure 12b, Section 3.4) and literature reports that gaps shorter than that are not perceivable in fluent speech (Nootheboom, 2007; Warren, 2013). The fact that this listener perceived this 83-ms-silence gap goes counter these findings reported in previous studies.

<sup>59</sup> He mentioned the pronunciation of *curled* distracted him and added that “because she make a pause ... before the, the word... and this pause can mean an emphasis or just she had to read it [to be more careful with the pronunciation]. Yeah, so, you can't say”. (Interview with F09, 00:09:36 – 00:10:00).

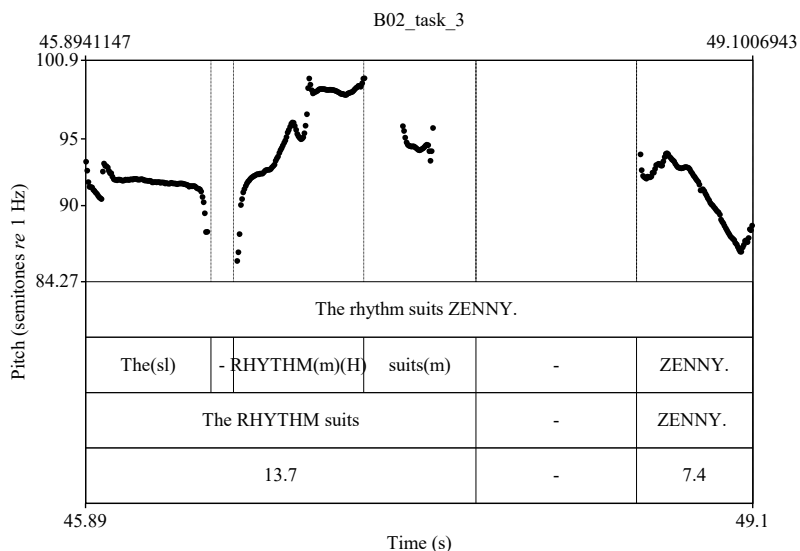


Figure 17: *No, the rhythm suits ZENNY* (Task 3, Item 03) by B02 – An example of NS placed in more than one location with a greater pitch change at an unexpected portion of the utterance.

The discursive context here is the same of that in Figure 15. Note that this production has two thought groups and two NSs: (1) *The RHYTHM suits* and (2) *ZENNY*. In the first thought group *the* has syllable lengthening, probably triggered by the challenge in pronouncing *rhythm* intelligibly, which ended up being pronounced [ˈrɪden]. In addition to that, *suits* was pronounced [swɪts]. It is not surprising that the NS in this thought group had a greater pitch change (13.7 semitones) than did the NS in the subsequent thought group (7.4 st). Additionally, these thought groups were separated by a long pause (771 ms). Even though the first thought group was more salient, *Zenny* was perceived as holding the NS and the expected interpretation was reached. F02 reported that he perceived more “energy” on *Zenny*. It may be that the speaker’s nod while saying *Zenny* may have helped the listener unconsciously decide for *Zenny*.

Tables 18 and 19 also display the percentages for the times the productions were plain, that is, they did not reach the three-semitone-pitch change, as, according to literature (t’Hart, 1981), a change smaller

than that alone may not be perceivable. Figure 18 illustrates this kind of production.

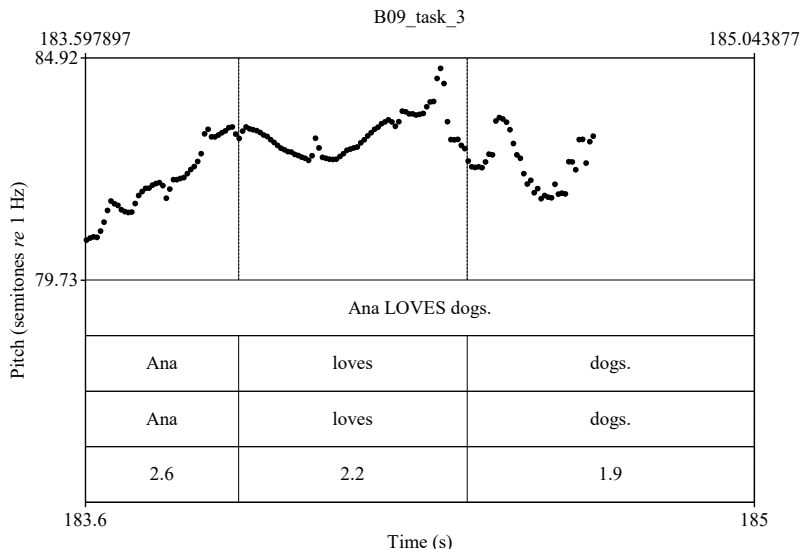


Figure 18: *No, Ana LOVES dogs* (Task 3, Item 07) by B09 – An example of utterances in which pitch changes did not reach the three-semitone-pitch change.

The discursive context for the production in Figure 10 was a decision being made on a pet to be given to Ana. B09 had to say the utterance in order to correct the information that Ana hated dogs. Therefore, she was supposed to place NS on *loves*. However, the figure reveals that there was not any perceivable pitch change to be called a NS as the greatest pitch change was that of 2.6 semitones in *Ana*, followed by *loves* with a pitch change of 2.2 semitones and *dogs* with 1.9 semitones. Despite the lack of prominence on *loves*, F09 perceived that portion of the utterance as the one having the NS. On the one hand, this appropriate perception would go counter the literature saying that variation smaller than three semitones is not perceptible (e.g., t'Hart, 1981). On the other hand, it may be argued that the appropriate perception here was probably because it was combined with a slight eyebrow movement, an extra visual cue as reported in the literature (Keating et al., 2003). However, it is hard to tell for sure if F09 could notice it or not or if it was just a guess.



This section approached the matter of NS position interfering or not with expected NS placement. It was found that having to place NS in initial, medial or final position was difficult for L1-BP speakers of English. Irrespective of positions, unexpected productions emerged with NS placements at an unexpected location or at more than two locations, or no NS at all was placed, as exemplified along this section. Despite all utterance positions have shown a high percentage of unexpected productions (77.8%, 81.7%, and 90.5% in initial, medial, and final positions, respectively), Hypothesis 1, that predicted that NS position would not play a role in the NS assignment, was not confirmed as a Chi-square test of independence showed that the difference in frequencies was significant for final position ( $\chi^2(1, N = 378) = 6.943, p = .0084$ ). This means that producing one NS only in final position was more challenging for these participants, as discussed in paragraph 7 of the this section. The section that follows is dedicated to understanding if type of information plays a role in the assignment of NS (RQ2).

#### **4.1.2 RQ2 – Production results according to type of information**

- RQ2: How is NS placed by L1-BP speakers of English, according to the type of information (being elicited, being corrected, being contrasted) in utterances?
- H2: Signaling eliciting information will pose more difficulties than contrasting and correcting information. These predictions are based on the assumption that contrasting and corrective information naturally require more prominence than providing eliciting information does (Frota & Moraes, 2016; Klok et al, 2011).

In this dataset, there were three types of information being conveyed by means of NS placement: information being elicited (Task 2), information being corrected (Task 3), and information being contrasted in questions (Task 4). Research Question 2 examined how NS is placed according to each type of information being conveyed. It was hypothesized (H2) that signaling information being elicited by means of NS assignment would be more challenging and thus would yield more unexpected productions than would signaling information being corrected and contrasted. This hypothesis comes from the fact that pitch change ranges in NS for correcting and contrasting information are wider than pitch change ranges for eliciting information (Frota &

Moraes, 2016; Klok et al, 2011) and thus may be easier to hear, which may facilitate production. Tables 20 and 21 display the results for NS according to the type of information being conveyed.

Table 20

*Production results according to type of information – first-time production*

Status	Elicited		Corrected		Contrasted		Total	
	N	%	N	%	N	%	N	%
Unexpected	96	76.2	103	81.7	116	92.1	315	83.5
Expected	30	23.8	23	18.3	10	7.9	63	16.6
Total	126	100.0	126	100.0	126	100.0	378	100.0

Table 21

*Production results according to type of information – repetition*

Status	Elicited		Corrected		Contrasted		Total	
	N	%	N	%	N	%	N	%
Unexpected	31	79.5	30	85.7	39	83.0	100	83.5
Expected	8	20.5	5	14.3	8	17.0	21	16.5
Total	39	100.0	35	100.0	47	100.0	121	100.0

In the first-time productions, the expected NS placement was more frequent when information was being elicited (23.8%, N = 30), followed by information being corrected (18.3%, N = 23) and finally by information being contrasted (7.9%, N = 10). Results were only slightly different for the repetition, in which information being elicited still yielded more expected NS placements (20.5%, N = 8), but this turn was followed by information being contrasted (17.0%, N = 8) and finally by information being corrected (14.3%, N = 5). Note, however, that information being contrasted yielded a greater number of repetitions (N = 47) probably because it was embedded in questions. A chi-square test of independence was performed to examine if these differences were statistically significant. The same procedures for doing the test described in Section 4.1.1 were adopted here. Tables 22 and 23 show the rearrangement of data and the partitioning of the contingency table in order to have two 2 x 2 tables instead of an only 2 x 3 table so the Chi-square test can be run.

Table 22

*Production results rearranged for statistical testing (2 x 3 Contingency Table) – IV Type of information*

Information being	Expected	Unexpected	
		N	N
Elicited		30	96
Corrected		23	103
Contrasted		10	116
Total		63	315

Table 23

*Production results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Type of Information*

Information being	Exp.	Unexp.	Information being	Exp.	Unexp.
	N	N		N	N
Elicited	30	96	Elicited + corrected	53	199
Corrected	23	103	Contrasted	10	116
Total	53	199	Total	63	315

The results of the chi-square for these two partitions<sup>60</sup> in Table 23 are that, for the leftmost table,  $X^2(1, N = 378) = 1.171, p = .2792$  and, for the rightmost table,  $X^2(1, N = 378) = 10.371, p = .0013$ . These results indicate that the difference lies between NS for information being contrasted when compared to elicited and corrected, but not between information being elicited and corrected.

As in the previous section, I opted for running multiple comparisons and thus I had the  $p$  value recalculated with Bonferroni correction ( $p < .01$ ). These multiple comparisons confirmed the ones reported in the previous paragraph. Differences were statistically significant only for information being contrasted if compared to information being elicited,  $X^2(2, N = 378) = 11.887, p = .0006$  and being corrected,  $X^2(2, N = 378) = 5.893, p = .0152$  but they were not significant between information being elicited and corrected,  $X^2(2, N = 378) = 1.171, p = .2792$ . This means that for these participants, placing NS as expected for contrasting information was more challenging than for eliciting or correcting information and this is statistically significant. Therefore, Hypothesis 2, which predicted that using NS to elicit

<sup>60</sup> Observe that with the partitioning the df for each of the partitions is 1 + 1, which sums to the 2 df for the full table.

information would be more challenging, was not confirmed, going counter this study's assumption that correcting and contrasting information would be easier to produce as expected.

The reasons for participants having more difficulties in placing NS to contrast information may lie on the fact that participants were making contrasts in questions. Asking questions has shown to be prosodically difficult for Brazilians (Passarella-Reis, 2014). Having to place NS in questions may have had a heavier role in this placement than the function played by the utterance: eliciting, contrasting or correcting.

Passarella-Reis (2014) found that Brazilians participating in her study tended to finish their polar questions<sup>61</sup> with a low boundary tone (L%), an intonational pattern that belongs to Brazilian Portuguese neutral polar questions, except for the region of Rio Grande do Sul (Frota et al., 2015). She also found that when attempting to reach the English intonational pattern for neutral polar questions, some participants produced a downstepped rise (!H), which means that they would rise intonation but the pitch change would not be great enough, a boundary tone also found among the productions of the questions in the dataset of the present study (N = 5). Her results signal the difficulty faced in terms of intonation when asking questions.

In the dataset of the present study, some participants seemed to be more concerned with rising their intonation to make their questions sound like a question than with placing NS as expected. Figure 14, in Section 4.1.1, provides an example of this instance. No NS was placed on the expected portion of the utterance, but only at the end to signal the question (*Does Zenny like SHRIMP?* instead of *Does Zenny LIKE shrimp?*). Even though there was this effort by the part of the speakers, only seven questions had rising intonation without a preceding pitch rising-falling movement and some had a low boundary tone (N = 28).

Most questions were asked with some pitch rising-falling movement in initial and/or medial position and pitch rising boundary (N = 86), which resembles a pattern described by Moraes (1998) – a double rise in BP non-neutral polar questions, one in the first stressed syllable, which contains a greater pitch change, and another one with a more reduced pitch change in the last stressed syllable. Questions with pitch changes in these locations in this dataset were considered unexpected productions, given that they do not match the English pattern (only one

---

<sup>61</sup> Frequently referred to as yes-no questions.

NS on the constituent holding the most important information and pitch compression on unimportant material). According to Moraes (1998), these pitch changes may signal (1) the beginning of a conversation, when the entire question is being introduced to the listener as new information, (2) rhetorical yes-no questions, (3) questions that begin by *I wonder/Do you think*, and (4) requests. None was the case of the questions comprising the dataset, but it may be the case that participants transferred these pitch movements from BP to English simply because they were questions.

This section introduced and discussed the production results for NS placements in relation to type of information. Hypothesis 2 was not confirmed, as utterances contrasting information had more unexpected NS placements than did the others. Next section addresses the effect of the presence of challenging words in terms of pronunciation.

#### **4.1.3 RQ3 – Production results according to complexity of words**

- RQ3: How is NS placed by Brazilian speakers of English, according to the complexity of words (non-challenging, short challenging, long challenging) in the utterances?
- H3: Based on the findings in the pilot study, Hypothesis 3 predicts that NS will be assigned as expected less frequently when in utterances with challenging words (either short or long) if compared to utterances with non-challenging words.

The results of the pilot study indicated that NS would sometimes have unexpected placements in utterances with challenging words in terms of pronunciation. Therefore, in the present study, the presence (or the absence) of challenging words was one of the independent variables investigated. Just for the reader to remember the utterances produced under each category – non-challenging, short-challenging, and long-challenging words – find in Table 24 all the utterances according to the complexity of words.

Table 24  
*Utterances produced in Tasks 2, 3, and 4.*

Status	Utterances	Position	Type of Information
Non-challenging	ZENNY loves you.	Initial	Elicited
	Zenny LOVES you.	Medial	Elicited
	Ana loves ZENNY.	Final	Elicited
	No, ANA sent the e-mail.	Initial	Corrected
	No, Ana LOVES dogs.	Medial	Corrected
	No, Zenny loves DOGS.	Final	Corrected
	Do MEN lie a lot?	Initial	Contrasted
	Does the meeting START at ten?	Medial	Contrasted
	Does the party start at TEN?	Final	Contrasted
Short challenging	ZENNY sent you the <b>pearls</b> .	Initial	Elicited
	The boss LOVES your <b>world</b> .	Medial	Elicited
	The T-shirt <b>shrank</b> A LOT.	Final	Elicited
	No, MARIA had it <b>curled</b> .	Initial	Corrected
	No, <b>pearls</b> MELT in <b>vinegar</b> .	Medial	Corrected
	No, the <b>rhythm suits</b> ZENNY.	Final	Corrected
	Is ZENNY afraid of <b>heights</b> ?	Initial	Contrasted
	Does Zenny LIKE <b>shrimp</b> ?	Medial	Contrasted
	Did Ana <b>twirl</b> HER HAIR?	Final	Contrasted
Long challenging	ZENNY <b>traveled</b> to <b>Pennsylvania</b> .	Initial	Elicited
	Zenny DISLIKES <b>vinegar</b> .	Medial	Elicited
	The <b>government</b> talked to ZENNY.	Final	Elicited
	No, RED is an <b>appetite</b> stimulator.	Initial	Corrected
	No, yellow STIMULATES <b>optimism</b> .	Medial	Corrected
	No, <b>unfortunately</b> Zenny AGREED.	Final	Corrected
	Is ZENNY leaving <b>immediately</b> ?	Initial	Contrasted
	Does this sequence REPRESENT the <b>hierarchy</b> ?	Medial	Contrasted
	Is the <b>atmosphere</b> around Zenny GOOD?	Final	Contrasted

Note: Uppercase letters: NS location; boldface: challenging words.

Hypothesis 3 predicted that NS would be assigned as expected less often if there were challenging words (either short or long) in the utterance than if there were no challenging words. Before going into the results for NS placement, the next paragraph introduces the speakers' impressions about their pronunciation of the words present in Tasks 2, 3, and 4, collected by means of the Pronunciation Self-Evaluation Test, in which they had to rate how well they could pronounce the words in a scale from 1 (one = I can not pronounce this word) to 9 (nine = my pronunciation is totally intelligible). Recall that the Pronunciation Self-Evaluation Test was used to gain insights about the speakers' view on whether the words included in the utterances were actually difficult as expected by the researcher and if this difficulty correlated with their production of NS. Table 25 displays the results of this test.

Table 25  
*Pronunciation Self-Evaluation Test - Speakers*

Speakers	Unintelligible		Maybe not intelligible		Intelligible enough		Totally intelligible	
	N	%	N	%	N	%	N	%
B01	5	10.2	2	4.1	10	20.4	32	65.3
B02	0		3	6.1	22	44.9	25	51.0
B03	0		3	6.1	28	57.1	18	36.7
B05	3	6.1	7	14.3	27	55.1	12	24.5
B06	0		4	8.2	17	34.7	28	57.1
B07	3	6.1	4	8.2	15	30.6	27	55.1
B08	1	2.0	5	10.2	17	34.7	26	53.1
B09	3	6.1	12	24.5	12	24.5	22	44.9
B10	2	4.1	8	16.3	32	65.3	7	14.3
B11	0		1	2.0	14	28.6	34	69.4
B12	0		10	20.4	36	73.5	3	6.1
B13	3	6.1	4	8.2	13	26.5	29	59.2
B14	2	4.1	2	4.1	4	8.2	41	83.7
B15	0		3	6.1	23	46.9	23	46.9
Mean	1.6	3.2	4.9	9.9	19.3	39.4	23.3	47.6

As can be seen (Table 25), the Pronunciation Self-Evaluation Test indicated that the speakers considered their pronunciation of 47.6% of the words as totally intelligible and 39.4% as intelligible enough, showing that they were at least comfortable about their pronunciation of

87% of the words (intelligible enough + totally intelligible). They were not confident about the pronunciation of 9.9% of the words and had no idea on how to pronounce 3.2% of the words in the tasks. The words mentioned in these two last categories (unintelligible and maybe not intelligible) varied among participants. They were: *curled*, *appetite*, *atmosphere*, *government*, *heights*, *hierarchy*, *immediately*, *leaving*, *melt*, *men*, *optimism*, *party*, *pearls*, *Pennsylvania*, *red*, *rhythm*, *sequence*, *shrank*, *shrimp*, *twirl*, *unfortunately*, *vinegar*, and *world*. It is interesting to note that, although most of the words mentioned were short- or long-challenging words, participants also mentioned words that were supposedly non-challenging words (e.g., *red* and *men*).

In order to examine if the way they rated their pronunciation of words related to their production of NS, a Pearson Correlation was run as the assumptions for this test were met. The Pearson Correlation test showed that there was no correlation between the reported pronunciation rate for each speaker and the total number of unexpected NS placements in the dataset ( $r = .026$ ,  $p = .930$ ).

The mean rate for each word was calculated in order to examine if these rates correlate with the total counts of expected productions. The ones considered the least intelligible were *curled* ( $M = 3.92$ ), *hierarchy* ( $M = 3.46$ ), *twirl* ( $M = 2.38$ ), and *vinegar* ( $M = 4.69$ ). The remainder of the words were considered intelligible enough ( $M > 5.0$ ) and very intelligible ( $M > 8.0$ ). As these data did not met the assumptions for a Pearson Correlation test, a nonparametric test was run. A Spearman's rho correlation showed that there is no correlation between the expected NS productions and the reported pronunciation of the words ( $r = .096$ ;  $p = .767$ ).

It is important to note that, although many words were rated as intelligible by the speakers, the listeners reported having difficulties to understand the speakers. This might mean that these speakers were not totally aware of how intelligible they sounded when pronouncing these words. However, an intelligibility test using a write-what-you-hear task would be necessary to support this assumption and doing that goes beyond the purposes of the present study. This is thus left for future investigation.

Having in mind the speakers' pronunciation ratings for their pronunciation of words in the tasks, one would expect that the pronunciation would not pose much of a problem for these speakers, given that they were more confident than they were insecure about the



pronunciation of the words. However, expected NS assignments tended to be more frequent in utterances with non-challenging words, as displayed in Tables 26 and 27.

Table 26

*Production results according to complexity of words – first-time production*

Status	Non-challenging		Short challenging		Long challenging		Total	
	N	%	N	%	N	%	N	%
	Unexpected	93	73.8	111	88.1	111	88.1	315
Expected	33	26.2	15	11.9	15	11.9	63	16.6
Total	126	100.0	126	100.0	126	100.0	378	100.0

Table 27

*Production results according to complexity of words – repetition*

Status	Non-challenging		Short challenging		Long challenging		Total	
	N	%	N	%	N	%	N	%
	Unexpected	27	75.0	32	88.9	41	83.7	100
Expected	9	25.0	4	11.1	8	16.3	21	16.5
Total	36	100.0	35	100.0	49	100.0	121	100.0

In the first-time productions, NS was more frequently assigned as expected in utterances with non-challenging words in terms of pronunciation (26.2%, N = 33) followed by utterances with short and long challenging words (11.9% each, N = 15).

In the repetitions, utterances with non-challenging words had NS assigned as expected 25% of the times (N = 9), followed by utterances with long challenging words (16.3%, N = 8) and then by those with short challenging words (11.1%, N = 4).

Some illustrations of these productions are provided in figures from 19 to 23<sup>62</sup>. Figure 13 (Section 4.1.1), repeated here for easy retrieval, illustrates an expected production of an utterance with non-challenging words.

<sup>62</sup> More illustrations are provided in Appendix T as well as a detailed discussion of the productions.

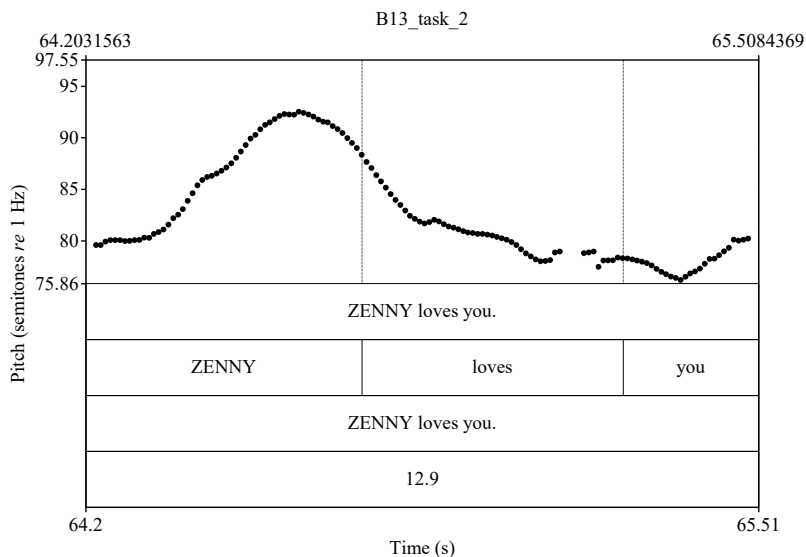


Figure 13: *ZENNY loves you* (Task 2, Item 04) by B13 – An example of an expected NS placement.

Figure 13 shows the production of *ZENNY loves you* by B13. Note that none of the words in the utterance should pose a challenge in terms of pronunciation for B13. The only word that could have had an expected pronunciation is *loves*, with vowel change or vowel epenthesis, but those were not the case. The easy words in terms of pronunciation here may have helped B13 to nuclear-stress the expected word, namely, *Zenny*, which was, in turn, perceived by F13, who grabbed the speaker's intent in this specific case. Figures 19 and 20 illustrate expected production in utterances with short-challenging and long-challenging words, respectively.

The context for the production in Figure 19 was that of business. People were talking about the project of a 'world' created by a coworker for a computer game and the speaker was expected to say the utterance as a response to the context question *How does the boss like my 'world'?* and thus, the word *loves* should be nuclear-stressed. The production was delivered with no pauses or pronunciation problems. Although some time was spent in saying *world*, *loves* had the pitch change of 4.8 semitones and more duration. The word *world* had many pronunciations in the dataset: [wɜrd], [wɜrdʒ], [wɜ:d], [ɔrd], and

[wɔʊd]. Even though B07 rated this word a 3, which means she was not sure if her pronunciation for this word was intelligible enough, the word did not pose any obstacles: it was pronounced as expected and the NS was also assigned as expected.

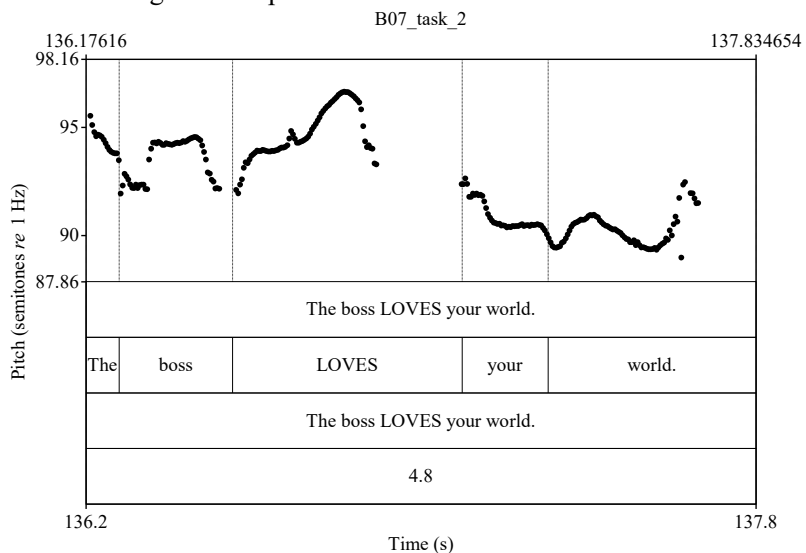


Figure 19: *The boss LOVES your world* (Task 2, Item 08) by B07 – An example of an expected production of utterances with short-challenging words.

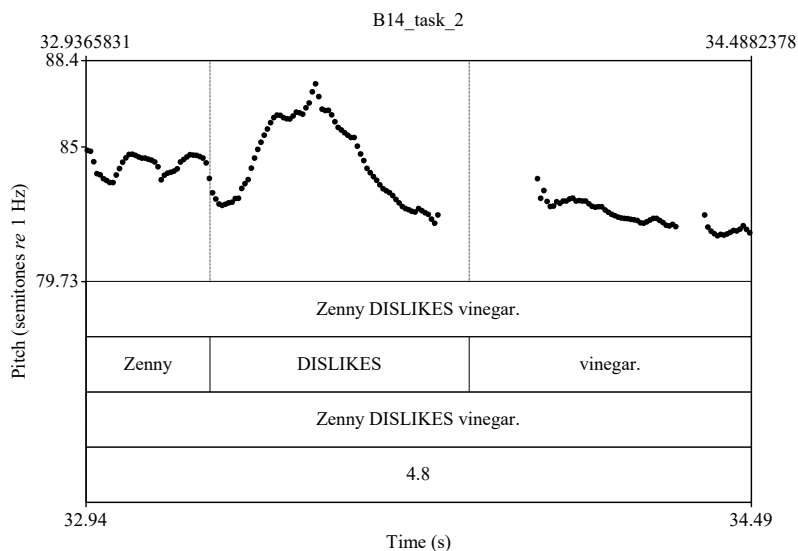


Figure 20: *Zenny DISLIKES vinegar* (Task 2, Item 03) by B14 – An example of an expected production of utterances with long-challenging words.

The context for the production in Figure 20 was the preparation of a special dinner for Zenny, the character created for data collection. B14 had to say the utterance in reply to the context question *How does Zenny like vinegar?* and thus the word *dislikes* should be nuclear-stressed. The challenging word here was *vinegar*, which was hardly ever pronounced as expected along the dataset, being often pronounced as [ˈvʌnɪgər] or [ˈvʌnəgər], and less often as [vɪˈnɛgər], [ˈvɪnɛgər], and [ˈvɪnɪgər]. B14 rated her pronunciation of *vinegar* a 9 (nine), which means she was very confident that the way she pronounced the word was intelligible. This confidence may have helped her with the accurate assignment of NS. Figures 21, 22, and 23 provide examples of unexpected productions of utterances with non-challenging, short-challenging, and long-challenging words.

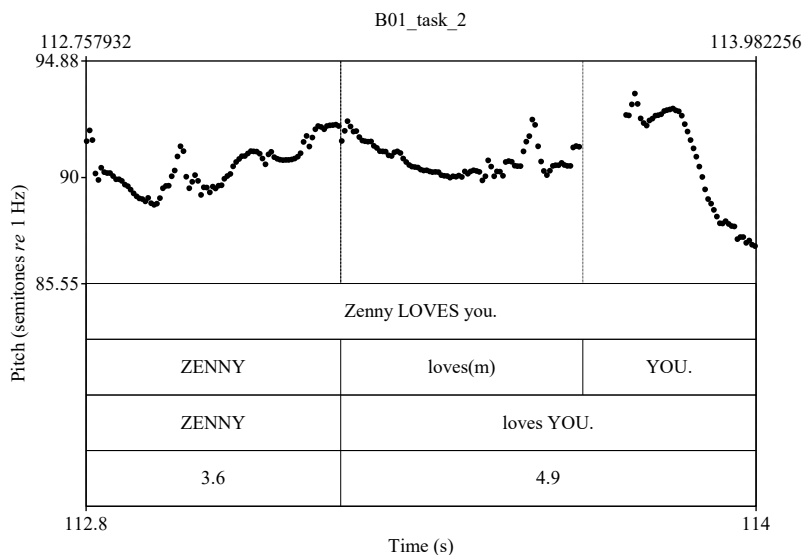


Figure 21: *Zenny LOVES you* (Task 2, Item 04) by B01 – An example of an unexpected production of utterances with non-challenging words.

The context for this utterance was first introduced in Figure 13 (produced by B13). The two productions are quite different. The pitch range here is not that great. *Zenny* has some prominence (a 3.6-semitone-pitch change) and *you* has the greatest pitch change (4.9 st). Emphasizing destressed words such as *you* is a common practice among Brazilians, as duly noted by Baptista (2001). The only word that should take NS here was *loves*, but it was destressed and it was pronounced as [ˈlɔvɪz]. This exemplifies the difficulty some participants had in placing NS in the middle of the sentence, going in line with Carpes (2014). As a result of unexpected NS placement, F01 had an unexpected interpretation for this production: that the question being answered was “Who does Zenny love?”.

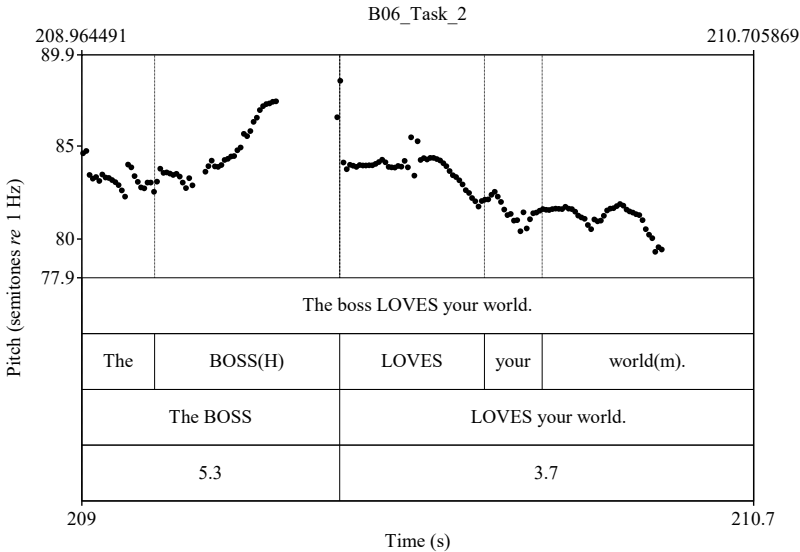


Figure 22: *The boss LOVES your world* (Task 2, Item 08) by B06 – An example of an unexpected production of utterances with short-challenging words.

The context for the utterance in Figure 22 was previously introduced in Figure 19. Observe that there is some rising intonation on *boss*, with a 5.3-semitone-pitch change. This gives the impression that there is more to be said but, at the same time, misguides the listener with regard to the NS. *Loves*, the expected portion for NS assignment, also has a pitch change, but smaller than the previous one: 3.7 st. *World* is mispronounced here in a way that the researcher was unable to phonetically represent it. Unexpected pronunciations in the dataset were [wɜrd], [wɜrdʒ], [wɜ:d], [ɔrd], and [wɔud]. Listening to this production once did not enable F06 to perceive the NS and interpret the utterance. Consequently, she asked B06 to repeat. It is my impression that when he repeated the utterance, B06 was more concerned with correcting his pronunciation than with placing the NS correctly, as he also nuclear-stressed *world*, pronounced [wɜr] in the repetition, with a 4.4-semitone-pitch change. F06 then misunderstood that they were talking about the object of the boss's love rather than about how the boss felt in relation to the *world* created for the computer game, an unexpected interpretation.

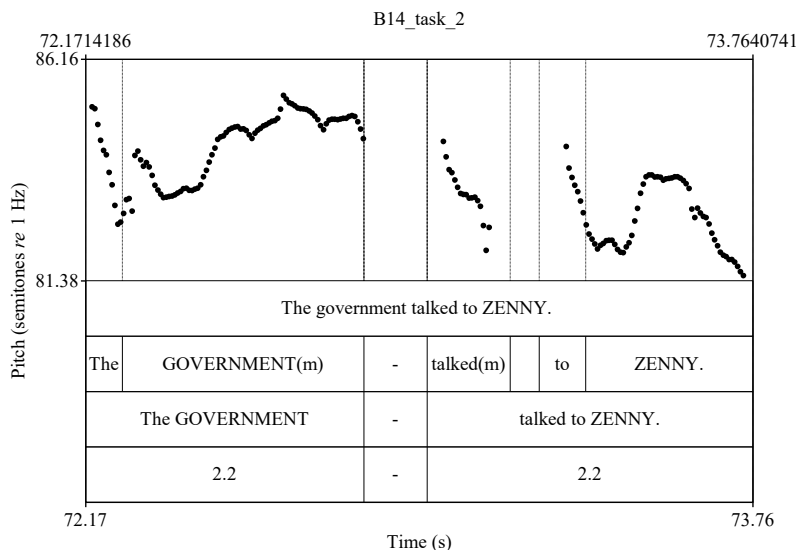


Figure 23: *The government talked to ZENNY* (Task 2, Item 06) by B14 – An example of an unexpected production of utterances with long-challenging words.

The context for this utterance was a visit of the government to the company. It was known that the government talked to someone specific and but it was not clear to whom. So, *Zenny* should be nuclear stressed. This specific production is particularly interesting as no portion of the utterance reaches a change up to three semitones. The two visible contours on *government* and *Zenny* are exactly the same: 2.2 st. The word *government* was produced as [ˈgʌvər̩mənt] and *talked* as [tɔk]. Other pronunciations found in the dataset were [ˈgʊvər̩mənt], [ˈgʊvər̩mənt], [ˈgʊvər̩nt], [ˈgʌvər̩mənt], [ˈgʊvər̩nəmənt], [ˈgʌrv̩mənt] and [tʌʊk], [ˈtɔkʰd], [ˈtʌʊkəd], [tʌʊkd], and [ˈtʌʊkɪd]. B14 thought that her pronunciation of *government* would probably be unintelligible, but that of *talked* would be completely intelligible. It is difficult to say what exactly happened here, but it is my impression that, being a quiet person, instead of getting loud on the word she was not sure about, B14

may have felt intimidated and used a lower tone for the whole sentence. Despite this, F14 could curiously get the expected message<sup>63</sup>.

Research question 3 inquires about the effect of the present of complex words in terms of pronunciation on the assignment of NS by BL1-BP speakers of English. Hypothesis 3 predicted that the presence of both short and long-challenging words would affect the performance of these speakers. The results displayed in Tables 26 and 27 show that unexpected NS placements were more frequent in utterances with short and long-challenging words than in utterances with non-challenging words. In order to check if the difference is statistically significant, the same procedures for chi-square testing performed in Sections 4.1.1 and 4.1.2 are taken here. Tables 28 and 29 show the rearrangement of data and the partitioning of the contingency table in order to have two 2 x 2 tables instead of an only 2 x 3 table so the Chi-square test can be run in the fashion suggested by Larson-Hall (2010).

Table 28

*Production results rearranged for statistical testing (2 x 3 Contingency Table) – IV Complexity of words*

Words	Expected	Unexpected
	N	N
Non-challenging	33	93
Short-challenging	15	111
Long-challenging	15	111
Total	63	315

Table 29

*Production results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Complexity of words*

Words	Exp.	Unexp.	Words	Exp.	Unexp.
	N	N		N	N
Non-challenging	33	93	Non-challenging + Short-challenging	48	204
Short-challenging	15	111	Long-challenging	15	111
Total	48	204	Total	63	315

<sup>63</sup> In this case, there was no way of comparing duration of the words, as they had quite different length, and no visual extra cues were used during the delivery of this utterance.



The results of the chi-square for these two partitions<sup>64</sup> in Table 29 are that, for the leftmost table,  $X^2(1, N = 378) = 8.338, p = .0039$  and, for the rightmost table,  $X^2(1, N = 378) = 3.086, p = .0790$ . These results indicate that the difference is statistically significant for short-challenging words only, which is awkward given that the frequencies are the same for short and long-challenging words.

As in the two previous sections, I opted for running multiple comparisons and thus I had the  $p$  value recalculated with Bonferroni correction ( $p < .01$ ). These multiple comparisons not only confirmed the significance for short-challenging words, but also showed that the difference was also statistically significant for long-challenging words when compared to non-challenging,  $X^2(2, N = 378) = 8.338, p = .0039$ . The multiple comparisons also showed that the difference is not statistically significant between short and long-challenging words,  $X^2(2, N = 378) = 0.000, p = 1.000$ . This means that for these participants, placing NS as expected in utterances containing non-challenging words was easier than in utterances containing challenging words, either short or long, as predicted in Hypothesis 3, that is, the presence of challenging words did affect the production of NS.

The difficulties posed by the presence of challenging words caused speakers to place NS in more than one location by producing relevant pitch changes on constituents before the hard word to pronounce or the word itself in an attempt to pronounce it as expected. Besides unexpected placements, unexpected pronunciations, unexpected pauses, unexpected rising intonation, unexpected syllable lengthening and repetition were also a result of the presence of this challenge in terms of pronunciation (see Appendix T). The impact of these aspects on speech processing may be great as will be discussed in Section 4.2.4.2.

In order to understand if utterances with challenging words had more of these aspects combined (making speech even more hard to process), the unexpected productions were grouped according to the number of aspects, the following results displayed in Table 30 emerge.

---

<sup>64</sup> Observe that with the partitioning the df for each of the partitions is  $1 + 1$ , which sums to the 2 df for the full table.

Table 30

*Number of aspects in unexpected productions: a comparison*

Status	Non-challenging		Short-Challenging		Long-Challenging		Total	
	N	%	N	%	N	%	N	%
None or only one aspect	60	64.5	48	43.2	30	27.0	138	43.8
2 or more aspects	33	35.5	63	56.8	81	73.0	177	56.2
Total	93	100.0	111	100.0	111	100.0	315	100.0

Table 30 summarizes the results for unexpected NS placements and the pronunciation aspects involved. It considers the results for unexpected productions that were free of problems (apart from unexpected NS placement) or had only one problem during its uttering, be it related to pronunciation of a word alone, or pause alone or syllable lengthening/rising intonation alone. The second row displays the sum of the results for unexpected productions that, besides the unexpected NS placement, included two or more aspects combined<sup>65</sup>.

These aspects were more frequent in utterances with long-challenging words (73%, N = 81), followed by those with short-challenging words (56.8%, N = 63) and then by those with non-challenging words (35.5%, N = 33). These results indicate that if obstacles in terms of pronunciation are present, the production gets noisier. F13 made an interesting metaphor that I call here the “*motorcycle metaphor*” and that I elaborate a little further. F13 argues that speaking a language other than one’s first is comparable to riding a motorcycle. When one is on the road, one has to (1) focus on the empty space and (2) drive through the obstacles (the other vehicles on the road, be them trucks, cars, or other motorcycles), by leaving them on one’s peripheral vision, in order to ride it successfully. If the rider focuses on

<sup>65</sup> The combinations that emerged from this dataset were: (1) unexpected pronunciation + unexpected pauses; (2) unexpected pronunciation + repetition; (3) unexpected pronunciation + unexpected pauses + repetition; (4) unexpected syllable lengthening/rising intonation + unexpected pauses; (5) unexpected syllable lengthening/rising intonation + unexpected pronunciation; (6) unexpected syllable lengthening/rising intonation + unexpected pronunciation + unexpected pauses; and (7) unexpected syllable lengthening/rising intonation + unexpected pronunciation + unexpected pauses + repetition. For examples of each combination, refer to Appendix T.

the obstacles, the rider crashes. When speaking another language, speaking the language itself is comparable to riding the motorcycle. If one focuses on the obstacles (here the difficult or challenging words), one sounds unintelligible and thus may not communicate successfully. Therefore, in order to get the message through, one has to focus on words that one knows<sup>66</sup>, otherwise, there will be breakdowns in communication. Section 4.2 introduces and discusses the interpretation results in order to try to understand whether these (expected or unexpected) productions caused breakdowns or yielded the expected interpretations.

#### **4.1.4 Summary of production results**

This section presented and discussed the production results according to location of NS, type of information, and level of complexity of words in terms of pronunciation of the utterances in Tasks 2, 3, and 4 (presented in Section 3.3), where information had to be elicited, corrected, and contrasted, respectively.

Production results showed that participants had difficulties to place NS as expected most of the times (Section 4.1). Hypothesis 1 predicted that there would be no differences in terms of placement of NS in the three positions investigated: initial, medial, and final and was not confirmed, given that utterances with NS in final position were more frequently placed as unexpected by the speakers (Section 4.1.1). Hypothesis 2 foretold that NS would be more frequently placed as expected in utterances correcting and contrasting information than in utterances eliciting information. This hypothesis was not confirmed as utterances contrasting information showed to be significantly more difficult for these speakers. It is important to note that an intervening variable appeared: utterances contrasting information were questions while utterances correcting and eliciting information were statements. Questions were way more difficult for them to produce since they involved the use of final rising intonation (Passarella-Reis, 2014; Moraes, 1998; Roelofsen & Gool, 2009). There was the tendency of rising in the beginning or middle of the sentence to fall and rise again at the end, which generated unexpected productions with more than one

---

<sup>66</sup> Focusing on known words and destressing hard words to pronounce was a strategy used by B14.

NS (Section 4.1.2). Hypothesis 3 predicted that expected assignments of NS would be more frequent in utterances with non-challenging words if compared to those with either short- or long-challenging words in terms of pronunciation. This hypothesis was confirmed.

Unexpected productions including more noise (unexpected pauses, unexpected pronunciation, repetition, and unexpected syllable lengthening/rising intonation) were more frequent in utterances with long- and short-challenging words, which means that facing utterances with these words made NS assignment and saying the utterances significantly more difficult for these speakers. The section that follows is dedicated to exploring how these productions were perceived and how the speakers' intent was interpreted.

#### 4.2 PERCEPTION OF NS AND INTERPRETATION OF BRAZILIANS' INTENT

We often find ourselves asking what someone is “trying to say.” And further along a scale of complexity and number of potential variables, we sometimes ask, or would like to be able to ask, “What did she mean *by* that?” (Nelson, 2008, p. 302)

This section presents and discusses the results of how Brazilians' intent was interpreted by 14 members of the international community that participated in the study as listeners: one English, two French, two German, two Italian, one Japanese, one Norwegian, one Polish, one Russian, one Slovak, one Chilean, and one Swiss.

These listeners engaged in four tasks with the Brazilians. In three of the tasks (Tasks 2, 3, and 4), they played the role of employees of an international company settled in Canada whose names were in a list of candidates to be fired. In order to keep their job, they had to prove that they had good listening skills, by listening to the speakers (Brazilians) and trying to understand the speakers' intent in what they were saying. They received a sheet containing three options that were possible choices for each utterance (Appendices K, M, and O). If they could not get what the speaker meant, they could ask him/her to repeat only twice. They would have their names crossed out of the firing list only if they succeed in the tasks. In general, they were very committed to the tasks,

got fully into their roles, and seemed very willing to converge with speakers' speech (Giles et al, 1991).

The listeners' answers for the questions in the tasks, their comments during the interviews, and their behavior or the need for repetition were considered in the analysis in order to map the utterances with successful interpretation of the speakers' intent, the sites of reduced interpretability and the real breakdowns in communication.

*Successful interpretations* were those that were done when heard only once and matched the expected interpretation according to the discursive context set for each utterance. *Sites of reduced interpretability* were considered as so when the listener resorted to repetition but the noise in communication did not render it unsuccessful, that is, the listener could grab the intended meaning by listening to the repetition and the speaker's intent was understood as expected. *Real communication breakdowns* refer here to two cases, which are described below:

1) when the listener asks for repetition and is still not able to understand what the speaker meant, and

2) when no repetition was asked and the listener thought he had understood what the speaker meant, but the listener was mistaken.

Table 31 displays the general results for the 364 utterances<sup>67</sup> according to the perception of NS.

Table 31  
*General perception results*

Status	N	%
Perceived as expected	162	44.5
Unexpected	202	55.5
Total	364	100.0

Table 31 shows that 44.5% (N = 162) of the utterances were perceived as expected while 55.5% (N = 202) were not. This means that unexpected productions (that represent 83.3% of the dataset) were not always misunderstood, as will be discussed in the following sections. Tables 32 and 33 show a closer look at all unexpected perceptions, in order to map the sites of reduced interpretability (only noise in communication) or real communication breakdowns.

<sup>67</sup> Recall that some cases were dismissed, as explained in Section 3.4.

Table 32

*Unexpected perceptions: Repetition meaning reduced interpretability only vs repetition leading to real communication breakdowns*

Unexpected Perception	Repeated Production	
	N	%
Breakdowns	42	36.8
Reduced Interpretability	72 <sup>a</sup>	63.2
Total	114	100.0

<sup>a</sup>. from which one interpretation was based on a third-time production.

Table 33

*Unexpected perceptions: real communication breakdowns – based on repetition Vs based on first production*

	Breakdowns	
	N	%
Repeated	42	32.3
Not repeated	88	67.7
Total	130	100.0

Tables 32 displays the percentages for real communication breakdowns (generated when the speakers intent was not interpreted as expected, despite the production being heard once, twice or three times) and sites of reduced interpretability (that required the repetition of the utterance in order to reach the expected interpretation of the speakers' intent). Table 33 displays the percentage of times that (1) real communication breakdowns happened even when negotiation of meaning happened (by means of repetition) and (2) they happened without negotiation of meaning (no repetition) meaning that neither the speaker nor the listener were aware that unsuccessful communication was taking place.

The figures in Table 32 show that 63.2% of the utterances that were repeated, that is, whose first production was not clear enough for the listener, were clear when heard for the second (N = 71) and the third (N = 1) times, rendering communication successful in these cases, despite the noise caused. The remaining 36.8% of the utterances were not clear and caused a real failure in communication (N = 42). A little more of concern is the fact that 67.7% of the real breakdowns in communication represent the utterances that were not repeated (N = 88), as displayed in Table 33. This means that most real communication

breakdowns happened when the listener was not aware of it. They assumed that they perceived the expected NS and interpreted the speakers' intent according to their perception. In real conversations, these could generate serious misunderstandings depending on the context of interaction. The next section discusses how expected and unexpected NS placements affected the interpretation of the speakers' intent.

#### 4.2.1 RQ4 – The interpretation of Speakers' intent – Expected vs unexpected

RQ4: How is the speakers' intent interpreted by IL-English listeners according to NS placement?

H4: Unexpected NS placement will hinder the interpretation of speakers' intent while expected NS placement will not (Atechi, 1994; Jenkins, 1997; Lanham, 1984; Tiffen, 1974).

Research question 4 inquired about how the speakers' intent would be interpreted by IL-English listeners according to NS placement. It was hypothesized that unexpected NS placements would yield unexpected interpretations while expected NS placements would yield expected interpretations. Recall that expected productions were those that had only one thought group and one NS placed at the expected location that corresponded to a given context. Unexpected productions were those in which NS was placed on an unexpected location or on more than one location in the utterance as defined in Section 3.4. Table 34 compares the frequencies for expected and unexpected productions in relation to their perception.

Table 34

*Production and perception: a comparison of frequencies*

	Expected Production		Unexpected Production		Total	
	N	%	N	%	N	%
Unexpected Perception	20	31.7	182	60.5	202	55.5
Expected Perception	43	68.3	119	39.5	162	44.5
Total	63	100.0	301	100.1	364	100.0

Table 34 shows that most unexpected NS assignments yielded unexpected NS perceptions (60.5%,  $N = 182$ ) and that expected NS placements were mostly perceived as expected (68.3%,  $N = 43$ ). A chi-square test of independence was performed to examine if these differences were statistically significant. The results of the chi-square was that  $X^2(1, N = 364) = 17.397, p < .0001$ , which indicates that the differences are statistically significant, confirming hypothesis 4.

Despite this statistical significance, the fact that there were mismatches between perception and production is undeniable. As can be seen from Table 34, 39.5% ( $N = 119$ ) of the unexpected NS placements were perceived as expected despite NS being misplaced, and 31.7% ( $N = 20$ ) of the expected NS placements were not perceived as expected. This means that communication was fragile between these participants. The listeners might have developed convergence strategies to accommodate to the speakers' speech (Giles et al., 1991) and these strategies might have worked for some items but not for all.

Listeners were told that they could ask the speaker to repeat every time they did not understand the speakers' intent interpreted according to the speakers' placement of NS. Because of that, when listeners needed, they had some utterances repeated once ( $N = 113$ ) or twice ( $N = 1$ ) or when they thought the message was clear, they did not have the utterances repeated ( $N = 250$ ). After the tasks, each listener was individually interviewed by the researcher. During the interview, they assessed their answers to the tasks, watched the video recordings of the tasks and were able to elaborate on some of the reasons why they asked for the speakers to repeat or why they did not perceive or why they perceived the NS. These results are displayed in Tables 35 and 36.



Table 35

*Listeners' opinions on why they interpreted the speakers' intent as expected.*

Reasons: Listener said that s/he	N	%
01 perceived as expected.	137	84.6
02 heard no distinct emphasis and thus relied on something else rather than on NS.	5	3.1
03 noticed more than one emphasis, but the expected portion was stronger.	7	4.3
04 noticed unexpected pauses.	1	.6
05 thought the pause was a cue to signal emphasis.	5	3.1
06 relied on the repetition of a word/phrase as to correct unexpected pronunciation or NS placement.	3	1.9
07 was attracted by unexpected pronunciation at the perceived portion.	3	1.9
08 had no idea of why s/he got it correct.	1	.6
Total	162	100.0

Table 36

*Listeners' opinions on why they did not interpret the speakers' intent as expected.*

Reasons: Listener said that s/he	N	%
01 perceived NS somewhere else.	46	22.8
02 perceived NS somewhere else and the perceived portion had unexpected pronunciation.	11	5.4
03 was distracted by unexpected pronunciation of a word in a portion other than the one with the perceived word.	9	4.5
04 relied on the repetition of a word/phrase as to correct unexpected pronunciation or NS placement.	5	2.5
05 heard no distinct emphasis and thus relied on something else rather than on NS.	40	19.8
06 noticed more than one emphasis.	31	15.3
07 noticed more than one emphasis and unexpected pauses.	2	1.0
08 was influenced by unexpected pauses and pronunciation issues.	1	.5
09 noticed unexpected pauses.	3	1.5
10 thought the pause was a cue to signal emphasis.	5	2.5
11 speaker was too fast and listener was still processing information.	3	1.5
12 doubted the speaker's production.	19	9.4
13 compared speaker's and listener's pronunciation and got distracted.	7	3.5
14 had no idea of why s/he got it wrong.	20	9.9
Total	202	100.0

As displayed on Table 35, eight explanations were given to the expected interpretations of the speakers' intent. According to the listeners, 84.6% (N = 137) of those which were interpreted as expected were so because NS was perceived as expected. Other reasons were related to the perception or not of emphasis: in (03), more than one emphasis was noticed, but the listeners perceived that the expected portion held the strongest (4.3%, N = 7), while in (02) no emphasis was noticed and listeners relied on something else rather than on NS to

interpret the speakers' intent (3.1%, N = 5) – duration, sentence structure, and sentence meaning. Listeners also thought that some pauses would be a cue to signal emphasis, as in (05), and noticed unexpected pauses but were still able to find the expected emphasis (explanation 04) (3.1% and .6%, respectively). Explanations (06) and (07) relate to unexpected pronunciation either by being attracted by it and thus choosing the unexpectedly pronounced word as the one holding the NS (1.9%, N = 3) or by the fact that the speaker repeated the portion with pronunciation issues (or mistaken NS) in order to correct it (1.9%, N = 3). Finally, only one instance of the expected interpretations was unexplained (in 08).

The explanations given for *not* interpreting the speakers' intent as expected were greater in number (14 explanations) (Table 36) than those displayed in Table 35 (8 explanations). Although more numerous, they still have to do with the ones discussed in Table 35, as they are related to the perception of none, one or more emphasis, pauses, and unexpected pronunciation. In the following paragraphs, I address these explanations and provide some examples about some reasons why the listeners failed to interpret the speakers' intents, namely, listener's disbelief about the speaker's ability to convey the message, unexpected pronunciation attracting the perception of NS, the lack of perception of distinctive emphasis, and the perception of too many pauses.

According to the listeners, they perceived NS on a position other than the expected in 22.8% (N = 46) of the unexpected perceptions (explanation 01). While 5.4% of the unexpected perceptions (N = 11) were perceived at an unexpected position which was also produced with unexpected pronunciation, 4.5% (N = 9) were perceived at an unexpected position, but in this case the unexpected pronunciation was in a portion of the utterance other than the perceived portion. In these two explanations (02 and 03), pronunciation interfered in two different ways: (a) by attracting the perception of NS on a unexpectedly pronounced word or (b) by distracting the listener from perceiving the NS on the expected word, because the unexpected pronunciation caused him/her not to hear the remainder of the utterance<sup>68</sup>. Figure 24 illustrates a production in which unexpected pronunciation distracted the listener from perceiving the NS as expected.

---

<sup>68</sup> An illustration of this case is provided in Section 4.2.4, Figures in T5 and 35.

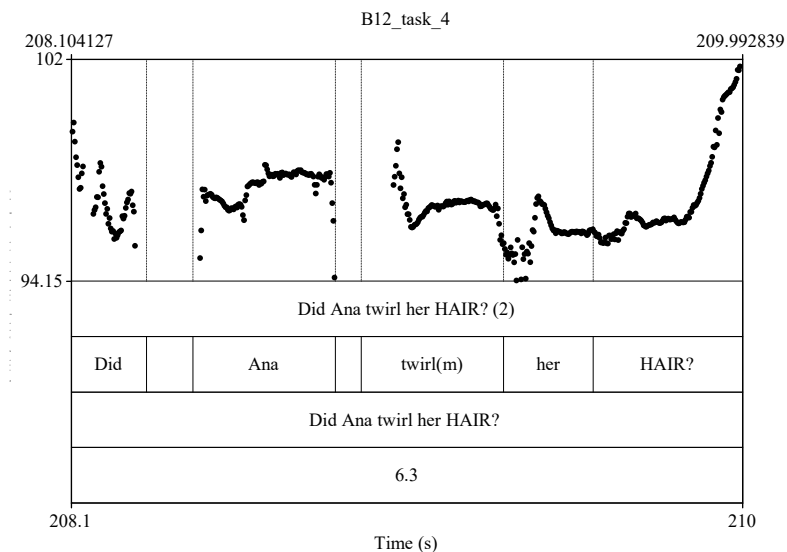


Figure 24: *Did Ana twirl her HAIR?* (Task 4, Item 07) repeated by B12 – An example of an unexpected perception caused by unexpected pronunciation as reported by the listener.

Figure 24 displays the repetition of *Did Ana twirl her HAIR* by B12. The expected position for NS was the word *hair*. The first time B12 asked the question, she did two equal pitch changes of 4.3 semitones in *twirl*, pronounced as [tʃu], and on *hair*, the expected location. Therefore, F12 asked her to repeat. In the second production, displayed in Figure 24, she managed to place NS only on *hair*. However, she changed the pronunciation of *twirl* into [twil]. As a matter of fact, this way of saying the word *twirl* affected F12's perception and prevented him from perceiving the NS on *hair*. F12 reported that he was confused by the mispronunciation of *twirl* and said he had to pay a lot of attention in order to process her speech and make sense of where the nuclear stress was placed:

I wasn't sure of what she was saying in the middle. [researcher says something] because I ahm, I'm trying hum, pick out which word she is putting emphasis on and then I'm confused about the word she's saying in the middle [Researcher: ahm, ok] I need a lot of attention, a lot of... I

think about it a lot. (Interview with F12, 00:44:45 – 00:45:15)

As a result, F12 interpreted that B12 wanted to know if Ana had twirled or tied her hair, an unexpected interpretation and a real communication breakdown.

Listeners reported that repetition in order to correct pronunciation or NS assignment misguided their decision 2.5% of the times (N = 5) (explanation 04)<sup>69</sup>. Listeners also reported that the reasons for unexpected interpretations were related to the perception of no distinct emphasis (19.8%, N = 40) or to the perception of more than one emphasis (15.3%, N = 31). In these cases, listeners relied on duration, sentence structure, sentence meaning or intuition (guesses) to opt for one interpretation. Figure 25 provides an example of these cases.

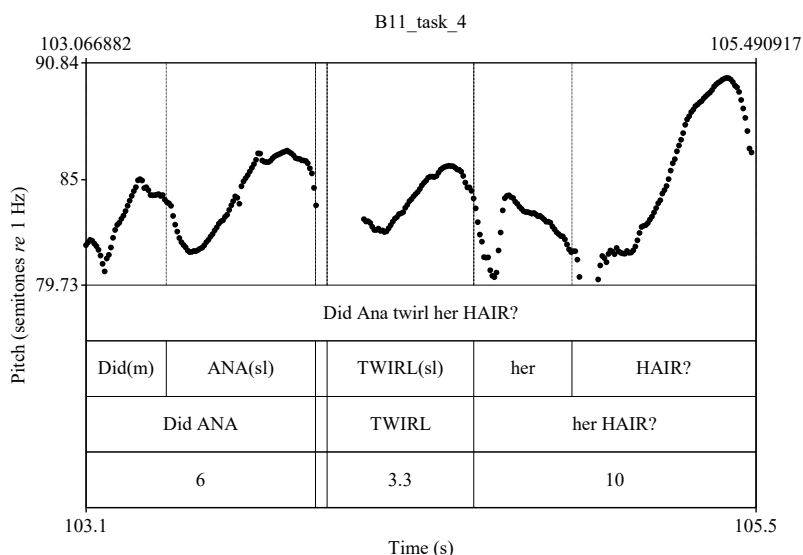


Figure 25: *Did Ana twirl her HAIR?* (Task 4, Item 07) by B11 – An example of listeners guiding their decisions on an aspect other than NS placement.

The reader is by now very acquainted with the utterance in Figure 25. The portion that should have been nuclear stressed is the one with

<sup>69</sup> An example is provided in Appendix T, by means of Figure T16.

the word *hair*. However, in this production, B11 did pitch changes on *Ana* (6 st, with syllable lengthening), *twirl* (3.3 st, with syllable lengthening), and *hair* (10 st). This number of inflections brought some confusion to the listener, who reported that she could not perceive any distinction between two of the places that were stressed: *Ana* and *hair*. Nevertheless, she did not ask B11 to repeat. She just preferred to consider the options and make a decision:

Maybe here I used the same thing. I didn't know so I, I didn't know if it was *Maria* or it was *her dress*. So I chose the third one. You know like I look at the possibilities and I choose which one is the most possible for me (...). My bad! I could've asked for him to repeat. (Interview with F11, Recording 2, 00:07:53 – 00:08:15)

Surprisingly, the analysis of the options led her to opt for the interpretation that *Ana* had twirled instead of tied her hair, which configured a real communication breakdown. Observe that *twirl* had the smallest pitch change, but still it was the choice made. This specific case in which the listener gives up on listening to what the speaker is saying is probably one of the greatest problems to communication in whatever context of interaction.

Explanations 07, 08, 09, and 10 are all related to pauses alone (1.5%, N = 3) or associated with another factor (4%, N = 8) and indicate that pausing is something noticeable and disturbing to some extent when it comes to perceiving NS and interpreting the speakers' intent, given that it is an important cue to thought group limits identification (Heusinger, 2007). Figure 26 provides one example of perceived pauses hindering communication.

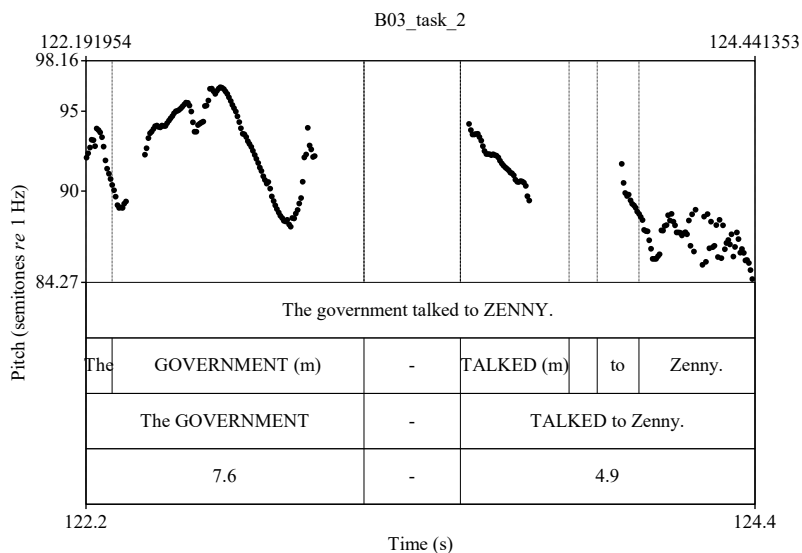


Figure 26: *The government talked to ZENNY* (Task 2, Item 06) by B03 – An example of perceived pauses hindering the expected interpretation of speakers' intent.

Figure 26 displays the Praat window for *The government talked to ZENNY*, uttered by B03, who was expected to place NS on *Zenny* in order to answer the question *Who did the government talk to?* B03 did two thought groups separated by a pause of 324 ms: (1) *The GOVERNMENT* (pronounced [*'gɒvər,mənt*]) and (2) *TALKED to Zenny* (*talked* was pronounced [*taʊk*]). Her voice came out creaky on *Zenny* and Praat could not give a fine tracking of the F0. However, by hearing the final part it is possible to assume that there was not a significant pitch change on *Zenny*, where the NS should fall. F03 reported that he could notice the emphasis in *government* and thought that the pause right after the word meant that *government* was the important information. Therefore, he misinterpreted that the question being answered was *Who talked to Zenny?*

Pauses were an interesting cue in this study, because listeners accommodated to the speakers' speech in different manners (Giles et al., 1991): for some listeners, a pause before a word meant that the *following word* would have the NS; for others, a pause after the word would mean that the *previous word* would have the NS; and there were

cases in which the same listener used both ways to relate pauses to NS locations. This is one more indication that supports the need to analyze these interactions individually in order to try to depict this complex and dynamic process that involves using language (Larsen-Freeman, 1997) and communicating.

Explanation (12) was unprecedented and happened 9.4% of the times (N = 19). It has to do with the fact that the listener doubted the speakers' production. It happened mainly in Task 4, because some speakers used rising intonation in questions, which caused a great pitch change at the end of the utterance, which coincided or not with the NS position. The listener would hear that pitch change and get in doubt if the speaker was really placing NS in that position or if the speaker was just thinking of making that utterance sound like a question, as exemplified in Figure 27.

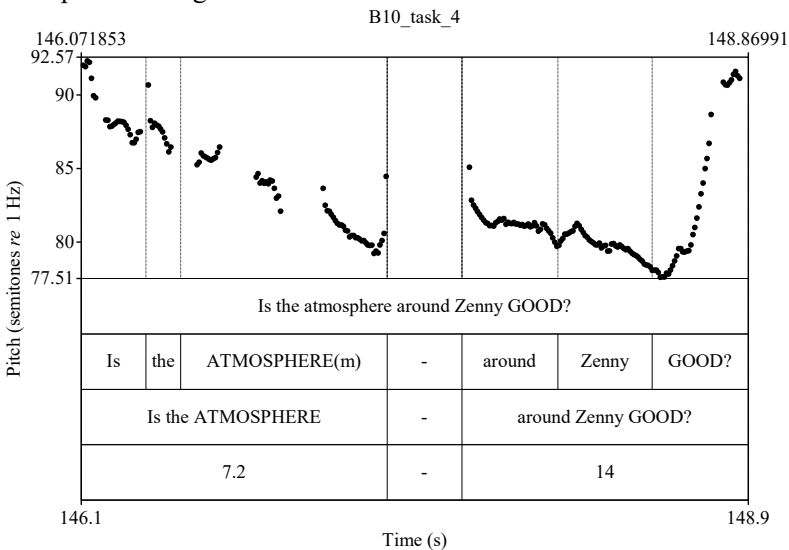


Figure 27: *Is the atmosphere around Zenny GOOD?* (Task 4, Item 05) by B10 – An example of listeners' uncertainty of final pitch change signaling important information or the pattern of a question.

The context for the utterance in Figure 27 is an informal conversation about the atmosphere at work when Zenny is around. The speaker had to ask the first half of the question *Is the atmosphere around Zenny good or bad?*, and thus had to nuclear stress the word *good* as if he were in doubt about the quality of the atmosphere when



Zenny was around. Figure 27 shows that B10 divided the utterance into two thought groups: (1) *Is the ATMOSPHERE* (pitch change of 7.2 st) and (2) *around Zenny GOOD* (pitch change of 14 st). Even though the pitch change was greater on *good*, F10 reported that she asked B10 to repeat because she could hear the emphasis on *good*, but she was suspicious that B10 nuclear-stressed *good* in an attempt to make his production sound like a question:

I think it was clear. It was hum ... good. But... hum... but I think that ... that ... hum... maybe he was trying to... he was doing the same thing again. He was just emphasizing the last word because it was a question... so... I wanted to be sure that... (Interview with F10, 00:05:25 – 00:05:52).

No mention was made to the unexpected pronunciation of *atmosphere* which sounded like [ˈætməʃ,fer]<sup>70</sup>. After she heard the repetition, F10 was able to perceive that the NS was on *good* and get the speaker's intent. Unfortunately, this behavior of doubting the speaker's production in interactions can cause some anxiety and cause discomfort during communication.

Listeners also reported that speakers' were too fast and they sometimes could not hear well what was being said as they were still processing the options (1.5%, N = 3). Finally 9.9% (N = 20) of unexpected interpretations were unexplained. The listener had no idea of why the intended message was not gotten.

These reasons in Tables 35 and 36 lead us to one conclusion: they all represent the listeners' attempt to accommodate to the speakers' speech. The hearing and processing of information is such a dynamic process that it involves many other processes and factors that are at play at the same time (Larsen-Freeman, 1997) that it is not feasible to give a sole reason for interpreting it properly or not. Besides, these processes are not always at a conscious level – such as the perception of visual

---

<sup>70</sup> Other unexpected pronunciation in this dataset for the word *atmosphere* were: [ˈʌtməs,fer], [əˈtʃiməs,fiər], [əˈtɪməs,fiər], [ˈætʃiməs,fiər], and [ˈætməs,feri].

cues<sup>71</sup> (Keating et al., 2003), and are not accessible to being described. The analysis shows that each listener has a different road to ride in order to converge and accommodate to the speaker's speech and understand the way the speaker's uses the language.

Hypothesis 4 predicted that the unexpected NS assignment would lead to communication problems (Atechi, 1994; Jenkins, 1997; Lanham, 1984; Tiffen, 1974), making the comprehension of the speaker's intent difficult (by causing sites of reduced interpretability) or impossible (real communication breakdowns) to grab. The hypothesis was confirmed. However, the present section discussed that the relation is not entirely straightforward, since accurate NS production not always led to NS accurate perception and consequent accurate intent interpretation as it (accurate perception) corresponds to (1) only 68.3% (N = 43) of the accurate productions, and to (2) 39.5% (N = 119) of the inaccurate productions in terms of NS (Table 32). This finding indicates that, when it comes to interpreting the speakers' intent, expected NS placement alone does not guarantee that communication will be successful since other intervening factors come into play, such as unexpected pronunciation, unexpected pauses and the convergence strategies used in order to accommodate to each other's speech (Giles et al., 1991). The next section addresses the effect of the expected NS location on the perception of NS and interpretation of speakers' intent.

#### **4.2.2 RQ5 – Perception results for the interpretation of speakers' intent according to NS position**

This section is divided into two parts. Section 4.2.2.1 presents the analysis of data in order to answer RQ5 which inquires about the influence of expected NS location on the perception of NS and the interpretation of speakers' intent. Section 4.2.2.2 presents the analysis of data regarding the perception of speakers' actual NS placement and the resulting interpretation of speakers' intent.

---

<sup>71</sup> None of the listeners reported perceiving a combination of emphasis (NS) with visual cues in order to interpret the speakers' intents. If this was the case, it was probably at an unconscious level.

### 4.2.2.1 Answering RQ5

- RQ5: How is the speakers' intent interpretation affected by the position of NS (initial, medial, final) in utterances?
- H5: Utterances with NS in final position will yield more unexpected perceptions than will those with NS in initial and medial position, because the final position is the canonical place for NS and narrow-focus NS has to increase greatly in order to be perceived in final position if compared to the other utterance positions (Calhoun, 2007).

Hypothesis 5 made predictions about the perception of NS according to NS position in the utterance. It was predicted that utterances with NS in final position would yield unexpected perceptions more often than would those in either initial or medial positions, due to the fact that, being the final the canonical NS position, narrow-focus NS in final position has to have a greater pitch change range (Calhoun, 2007). This means that even pitch changes with smaller pitch range (but greater than 3 st) in initial and medial positions are easier to perceive. Table 37 displays the perception results according to NS position.

Table 37

*Perception results according to expected NS position*

Status	Initial		Medial		Final		Total	
	N	%	N	%	N	%	N	%
Expected	65	52.8	47	39.2	50	41.3	162	44.5
Unexpected	58	47.2	73	60.8	71	58.7	202	55.5
Total	123	100.0	122	100.0	121	100.0	364	100.0

The results in Table 37 show that utterances with NS in initial position were perceived as expected more often (52.8%, N = 65), followed by those with NS in final position (41.3%, N = 50), and then by those with NS in medial position (39.2%, N = 47). A chi-square test of independence was performed to examine if these differences were statistically significant. The same procedures for doing the test described in Section 4.1.1 were used here. Tables 38 and 39 show the rearrangement of data and the partitioning of the contingency table in order to have two 2 x 2 tables instead of an only 2 x 3 table so the Chi-square test can be run (Larson-Hall, 2010).

Table 38

*Perception results rearranged for statistical testing (2 x 3 Contingency Table) – IV NS Location*

Location	Expected		Unexpected	
	N		N	
Initial	65		58	
Medial	47		73	
Final	50		71	
Total	162		202	

Table 39

*Perception results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV NS Location*

Location	Exp.	Unexp.	Location	Exp.	Unexp.
	N	N		N	N
Initial	65	58	Initial + Medial	112	131
Medial	47	73	Final	50	71
Total	112	131	Total	162	202

The results of the chi-square for these two partitions in Table 39 are that, for the leftmost table,  $X^2(1, N = 364) = 4.574, p = .0325$  and, for the rightmost table,  $X^2(1, N = 364) = 0.744, p = .3885$ . These results indicate that perception of NS in initial position is easier than in medial and final positions and the difference is statistically significant.

Using the same reasoning of the analysis of production data, I opted for running multiple comparisons and thus I had the  $p$  value recalculated with Bonferroni correction ( $p < .01$ ). These multiple comparisons did not confirm the statistical difference for initial position in relation to medial position,  $X^2(2, N = 364) = 4.574, p = .0325$  (significant at .01), and final position,  $X^2(2, N = 364) = 3.250, p = .0714$ . However, they confirmed the lack of statistical significance for the difference between medial and final positions,  $X^2(2, N = 364) = 0.116, p = .7330$ . Therefore, Hypothesis 5 was disconfirmed, as unexpected perception was not more frequent for NS in the end of utterances. These results indicate that the perception of NS and the consequent interpretation of speakers' intents were not connected to the expected place of NS in the utterance.

This section looked into the perception of NS and the interpretation of speakers' intents in relation to the expected placement of NS in utterances. The next section approaches perception by the listeners in relation to the actual NS placements by the speakers.

#### 4.2.2.2 Perception results according to actual NS locations as placed by L1-BP speakers

In order to understand the extent to which the productions were perceived as expected according to the actual location of NS, a comparison between production and perception was drawn, given that speakers did not place NS on or only on the expected position most of the times (83.3% when produced for the first time and 83.5% when repeated, as seen in Table 13). Table 40 displays the results of such a comparison.

Table 40

*Perception of expected and unexpected productions - Location*

NS Location	Production	Perception	N	%
Initial	Unexpected	Unexpected	50	52.6
		Perceived as expected	45	47.4
		Total	95	100.0
	As expected	Unexpected	8	28.6
		Perceived as expected	20	71.4
		Total	28	100.0
Medial	Unexpected	Unexpected	65	67.0
		Perceived as expected	32	33.0
		Total	97	100.0
	As expected	Unexpected	8	34.8
		Perceived as expected	15	65.2
		Total	23	100.0

(Continued)

(Table 40 continued)

NS Location	Production	Perception	N	%
Final	Unexpected	Unexpected	67	61.5
		Perceived as expected	42	38.5
		Total	109	100.0
As expected	As expected	Unexpected	4	33.3
		Perceived as expected	8	66.7
		Total	12	100.0

The figures in Table 40 shows that expected productions yielded expected interpretations more frequently when NS was in initial position (71.4%, N = 20), followed by utterances with NS in final position (66.7%, N = 8) and then by those with NS in medial position (65.2%, N = 15). Therefore, the unexpected productions rendered unexpected interpretations more often for utterances with NS in medial position (67.0%), followed by those with NS in final position (61.5%), and then by those with NS in initial position (52.6%).

Note in Table 40 that a little less than half of the unexpected productions enabled the listener to still interpret as expected the speakers' intent in initial position (47.4%). In order to understand the mismatch between production and perception according to NS position, the frequencies for production and perception were run taking into account the production and perception details gathered through the analysis of the utterances and the answers collected from the listeners' sheets during tasks 3, 4, and 5. These results are displayed in Tables 41, 42, and 43.

Table 41

*Perception and production details according to NS position in the utterances – expected allocations*

NS Location	Production	Perception	N	%
Initial	As expected	As expected	20	71.4
		Perceived as medial	1	3.6
		Perceived as final	1	3.6
		As expected after repeated	2	7.1
		Listener doubted NS place	4	14.3
		Total	28	100.0
Medial	As expected	As expected	15	65.2
		Perceived as final	1	4.3
		As expected after repeated	4	17.4
		Listener doubted NS place	3	13.0
		Total	23	100.0
Final	As expected	As expected	8	66.7
		Perceived as medial	1	8.3
		As expected after repeated	1	8.3
		Listener doubted NS place	2	16.7
		Total	12	100.0

Table 41 displays the way the expected productions were interpreted according to the position of NS in the utterances. Irrespective of positions, when misinterpreted, NS was perceived in another location, or was gotten after repetition, either because the listener had not gotten it when he or she heard it for the first time or because the listener doubted that the speaker had placed NS as expected. Therefore, the unexpected perceptions were mostly only sites of reduced interpretability that were cleared out after the utterances were heard for the second time. This means that expected productions usually led to successful communication, but some had some turbulence on the way. Figures 28 and 29 provide an example of an expected production that had some noise but ended up being interpreted as expected.

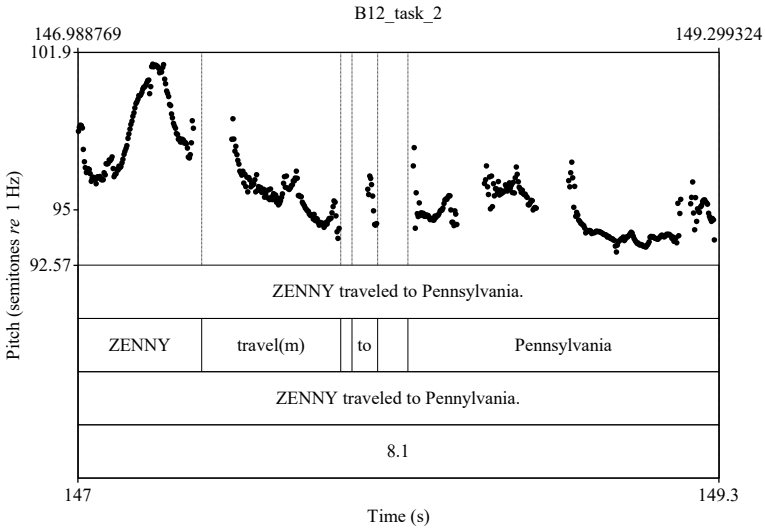


Figure 28: *ZENNY traveled to Pennsylvania* (Task 2, Item 03) by B12 – An example of an expected NS placement that had sites of reduced interpretability, but was interpreted as expected (first-time production).

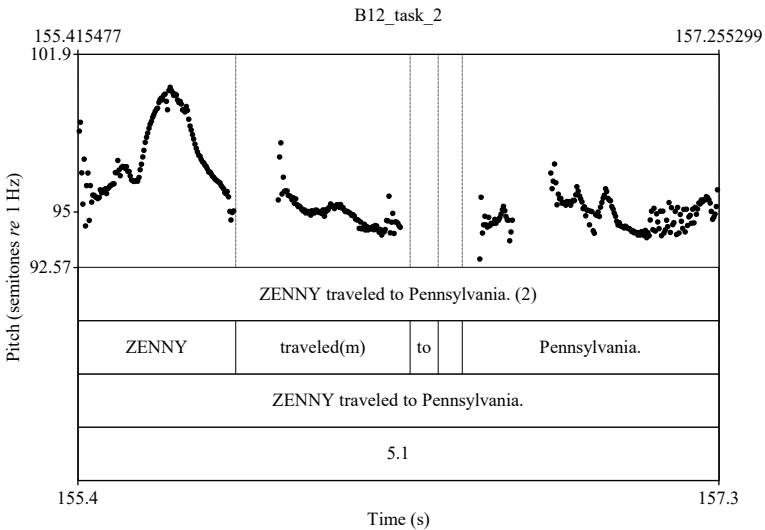


Figure 29: *ZENNY traveled to Pennsylvania* (Task 2, Item 03) by B12 – An example of an expected NS placement that had sites of reduced interpretability, but was interpreted as expected (repetition).



The context for the utterance in Figures 28 and 29 is a conversation in which an airplane ticket had to be bought for someone from Pennsylvania back to Canada. The listener has to issue the ticket but the listener does not know who traveled to Pennsylvania. B12 had to say the utterance by eliciting that the person who traveled to Pennsylvania was Zenny, by placing NS on *Zenny*. During Task 2, B12 tended to place some pitch change at the beginning of sentences. This pitch change was held sometimes until the middle portion of the utterances. When she meant to emphasize the middle, the pitch change was not that great in the medial location<sup>72</sup>. In *ZENNY traveled to Pennsylvania*, the expected position for NS is already initial. Therefore, B12 managed to place NS as expected, as can be seen in Figures 28 and 29. However, F12 seemed to have already converged and accommodated to B12's pitch changing in the beginning of the utterance and was not sure whether what he perceived was the NS or just B12's tendency of initial-pitch changing. Consequently, he asked her to repeat. In the repetition, the NS was again on *Zenny*. Then, F12 assumed that the initial position was the position that she meant to emphasize and interpreted her intent as expected.

Table 42 displays the frequencies for unexpected productions when they had NS placed in a portion of the utterance other than the expected portion as discussed in Section 4.1.1.

---

<sup>72</sup> Her difficult matches that of other L1-BP speakers in focusing constituents in the middle portion of utterances, as reported by Carpes (2014).

Table 42

*Perception and production details according to NS position in the utterances – unexpected allocations (Case 01)*

NS Location	Production	Perception	N	%	
Initial	As medial	As expected	1	50.0	
		As expected after repeated	1	50.0	
		Total	2	100.0	
	As final	Perceived as final	1	100.0	
		Total	1	100.0	
	Medial	As initial	Perceived as initial	1	50.0
Perceived as final			1	50.0	
Total			2	100.0	
As final		Perceived as final	1	50.0	
		As expected after repeated	1	50.0	
		Total	2	100.0	
Final		As initial	Perceived as expected	1	20.0
			Perceived as initial	1	20.0
			Perceived as medial	1	20.0
	As expected after repeated		2	40.0	
	Total		5	100.0	
	As medial	Perceived as expected	1	33.3	
		Perceived as medial	2	66.7	
		Total	3	100.0	

Unexpected productions with NS placed in a portion of the utterance other than the expected portion were few (N = 14). The frequencies show that a misplaced NS was not always perceived at the misplaced position. For instance, a NS placed in initial position that should have been placed in final position was perceived as expected (as final). That happened for misplacements in initial position too, but not for those in medial position: Every time a misallocation of a NS that should be in medial position happened, reduced interpretability (N = 1) or communication breakdowns took place (N = 3). This may corroborate the indication that there is a certain preference for listeners to perceive NS in initial and final positions. Figure 30 exemplifies the production of an utterance with NS in final position misplaced to initial position and still being perceived in final position (as expected).

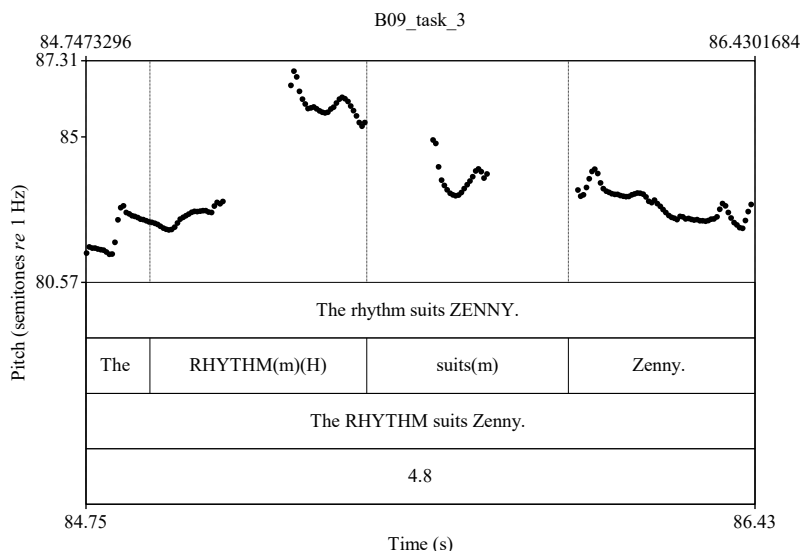


Figure 30: *The rhythm suits ZENNY* (Task 3, Item 03) by B09 – An example of a NS assigned to an unexpected position perceived at the expected position.

The utterance displayed in Figure 30 belongs to Task 3, in which speakers corrected mistaken information. B09 had to nuclear-stress *Zenny* in order to correct that the rhythm suited him, not Ana. In Task 3, B09 tended to produce pitch changes in the beginning of the utterances, sometimes combined with the expected location. Figure 30 shows that for this utterance the assignment of the NS was only placed on an unexpected location, namely, on *rhythm*, which was pronounced [ˈrɪtʃəm], with rising intonation. Even though no perceivable NS was placed on *Zenny*, the expected location, F09 interpreted the utterance as expected. It may be that F09 developed a convergence strategy (Giles et al., 1991) to accommodate to B09’s tendency of producing pitch changes in the beginning of the sentences and ignored those changes. As the video recording shows, *Zenny* was produced in tandem with an eyebrow movement, a visual cue for NS placement (Keating et al., 2003). However, it is not possible to say that F09 saw the eyebrow movement. F09 reported that he thought that there was too much emphasis and that this utterance sounded unnatural, but he did not say why he chose the final as the position having the NS.

Table 43 displays the frequencies for the production in which no portions of the utterance had a three-semitone-pitch change (case 02), NS was placed in two or more locations and the expected location for NS held the greatest pitch change (case 03), NS was placed in two or more locations with no differences in pitch changes (case 04), and NS was placed in two or more locations and the greatest pitch change was placed at an unexpected portion of the utterance (case 05).

Table 43

*Perception and production details according to NS position in the utterances – placement of NS in two or more locations or nowhere – Cases 02, 03, 04, and 05*

NS Location	Production	Perception	N	%
Initial (2)	Nowhere	Perceived as final	1	100.0
		Total	1	100.0
(3)	2 or more locations: greater on the expected	Perceived as expected	27	69.2
		Perceived as medial	1	2.6
		Perceived as final	4	10.3
		As expected after repeated	5	12.9
		Listener doubted NS place	2	5.1
		Total	39	100.0
(4)	2 or more locations: no differences	Perceived as expected	9	31.0
		Perceived as medial	7	24.1
		Perceived as final	8	27.6
		As expected after repeated	3	10.3
		Listener doubted NS place	2	6.9
		Total	29	100.0
(5)	2 or more locations: greater pitch change at an unexpected location	Perceived as expected	8	34.8
		Perceived as medial	3	13.0
		Perceived as final	11	47.8
		Listener doubted NS place	1	4.3
		Total	23	100.0

(continued)

(Table 43 continued)

NS Location	Production	Perception	N	%
Medial	(2) Nowhere	Perceived as expected	5	83.3
		Perceived as final	1	16.7
		Total	6	100.0
	(3) 2 or more locations: greater on the expected	As expected	10	55.6
		Perceived as initial	2	11.1
		Perceived as final	2	11.1
		As expected after repeated	4	22.2
		Total	18	100.0
	(4) 2 or more locations: no differences	Perceived as expected	11	26.2
		Perceived as initial	10	23.8
		Perceived as final	10	23.8
		As expected after repeated	9	21.4
		Listener doubted NS place	2	4.8
	Total	42	100.0	
	(5) 2 or more locations: greater at an unexpected location	Perceived as expected	6	22.2
Perceived as initial		8	29.6	
Perceived as final		9	33.3	
As expected after repeated		4	14.8	
Total		27	100.0	
Final	(2) Nowhere	Perceived as expected	4	100.0
		Total	4	100.0
	(3) 2 or more locations: greater on the expected	Perceived as expected	12	40.0
		Perceived as initial	7	23.3
		Perceived as medial	4	13.3
		As expected after repeated	6	20.0
		Listener doubted NS place	1	3.3
	Total	30	100.0	
	(4) 2 or more locations: no differences	Perceived as expected	16	38.1
		Perceived as initial	13	31.0
		Perceived as medial	6	14.3
		As expected after repeated	7	16.7
	Total	42	100.0	
	(5) 2 or more locations: greater pitch at an unexpected location	Perceived as expected	8	32.0
		Perceived as initial	8	32.0
Perceived as medial		3	12.0	
As expected after repeated		5	20.0	
Listener doubted NS place		1	4.0	
Total	25	100.0		

When a three-semitone-pitch change was not produced (case 2), NS was always perceived as expected for final position (100.0%, N = 4), most of the times for medial position (83.3%, N = 5) and perceived as final in initial position. This indicates a preference for final position when there is no distinct pitch change, possibly because of NS canonical final position, that is, the absence of a three-semitone-or-greater-pitch change does not harm communication at all when NS is at the end of the utterance. Figure 23, used in Section 4.1.3 and in Appendix T as Figure T12, reproduced here for the readers' comfort, provides a fine illustration of utterances that were produced with no NS and perceived as expected.

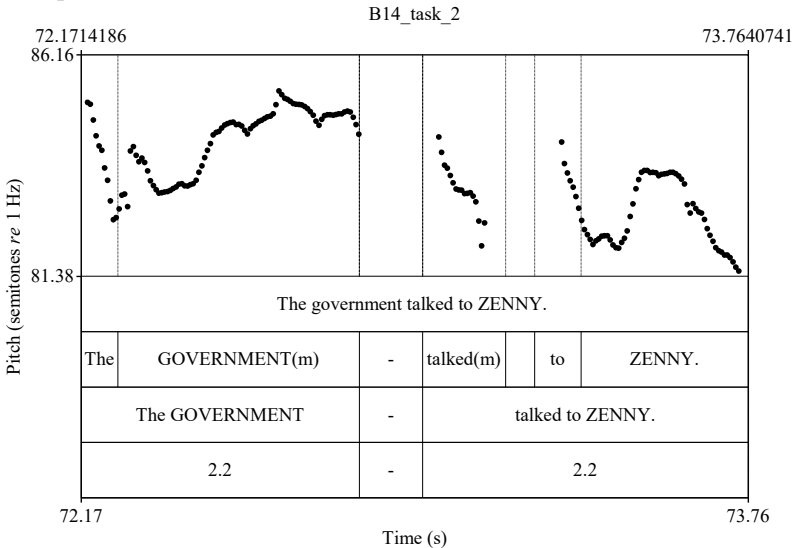


Figure 23: *The government talked to ZENNY* (Task 2, Item 06) by B14 – An example of an unexpected production of utterances with long-challenging words.

Recall that the production in Figure 23 had two small equal pitch changes of 2.2 semitones, one on *government* and another one on *Zenny*, and there were no extra cues combined with the production. However, F14 was still able to grab the expected interpretation that the person to whom the government had talked was Zenny. It is possible that F14

chose this interpretation because, as no NS could be perceived<sup>73</sup>, he opted for the NS canonical final position.

When an utterance was said with two or more NSs and the one at the expected portion of the utterance held the greatest pitch change (case 3), perception was as expected most of the times for utterances with NS in initial position (69.2%, N = 27), followed by those with NS in medial position (55.6%, N = 10) and then by those with NS in final position (40 %, N = 12). The latter had NS perceived as initial 23.3% of the times (N = 7), even though the NS assigned on the final position held the greatest pitch change. It is possible that, for these instances, the listener assumed that the speakers would produce final pitch changes to signal questions rather than NS. This kind of production hindered communication to a greater extent when NS was expected to fall in the end of the utterances. Figure 31 illustrates this occurrence.

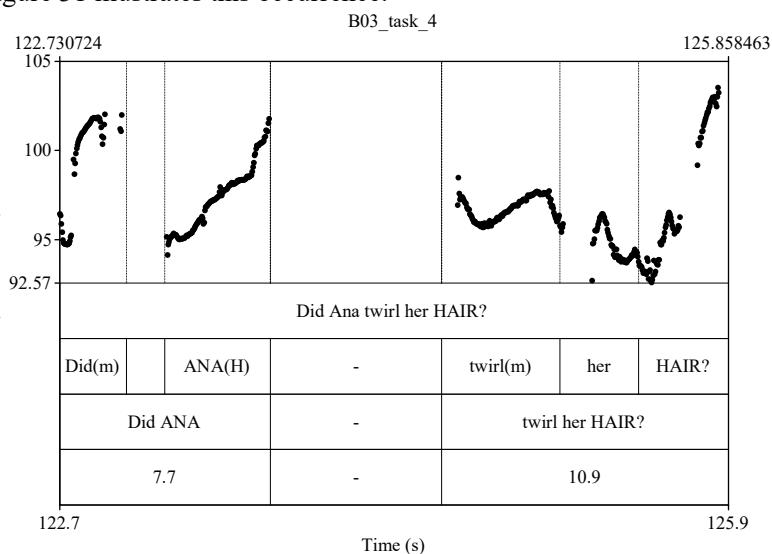


Figure 31: *Did Ana twirl her HAIR?* (Task 4, Item 07) by B03 – An example of a communication breakdown resulted from an utterance said with two or more NSs, with a greater pitch change range on the expected portion.

In the utterance in Figure 31, B03 had to nuclear-stress *hair*. In Task 4, her pitch changes were great and unexpected pauses were

<sup>73</sup> Unfortunately, this production was not thoroughly discussed during the interview.

frequent. The figure shows the production of two thought groups with great pitch changes, separated by a pause (802 ms) preceded by rising intonation: (1) *Did ANA* (7.7 st) and *twirl her HAIR* (10.9 st), the latter having the greatest pitch change. Even though the expected portion had the greatest pitch change, F03 probably considered that the pitch change on *hair* was a result of the rising intonation expected for questions and chose the first thought group as the one with the NS holding the important information, which rendered communication unsuccessful.

When the utterance was said with two or more NSs with no differences in pitch changes (case 4), NS was perceived as expected more often in final position (38.1%, N = 16), followed by those in initial position (31.0%, N = 9), and then by those in medial position (26.2%, N = 11). Therefore, it seems that ‘no distinct pitch change ranges’ do not harm NS perception in final position as much as in initial and medial positions, as the listener is still able to find the emphasis, probably because of NS canonical position (Liberman, 1975). An example of these perceptions is given in Figure 32.

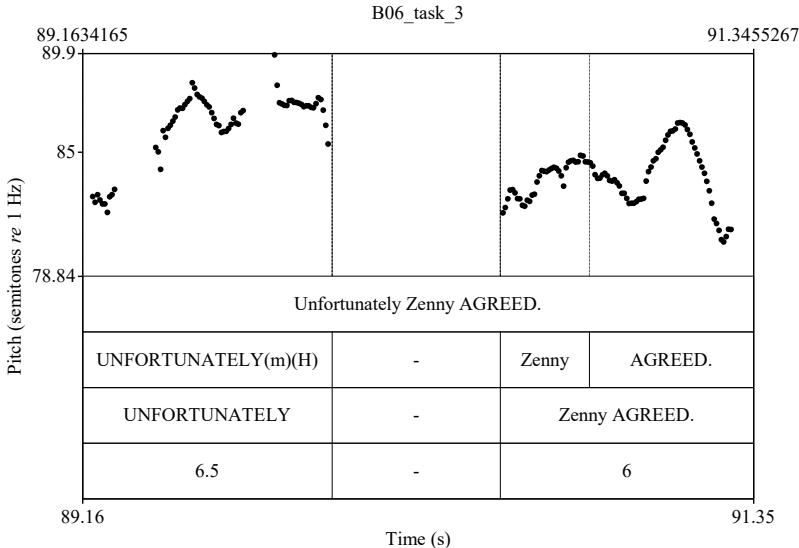


Figure 32: *Unfortunately Zenny AGREED* (Task 3, Item 05) by B06 – An example of an expected perception of NS in an utterance said with two or more NSs, with no distinct pitch change ranges.



In the utterance in Figure 32, NS should be placed on *agreed* in order to correct the information that Zenny had disagreed to move to Brazil. In Task 3, B06 had some greater pitch changes sometimes accompanied by smaller ones. Usually, when the NS was expected to be placed in initial position, pitch change was greater in the beginning of the sentence and smaller at the end. Maybe it was so due to the L% to end up the sentence. Note, however, that when the expected position was the final, as the one in Figure 32, some pitch change would also take place at the beginning of the sentence. Figure 32 shows then two locations for NS, one in each of the two thought groups produced: (1) *Unfortunately* (6.5 st), pronounced [ʌn'fɔrtʃʊ,nɛtli] with rising intonation, and (2) *Zenny AGREED* (6 st). These were separated by a 547-millisecond pause. Even though the two NSs had similar pitch changes, F06 ignored the first and understood that the final portion of the utterance held the important information and succeeded in interpreting the speaker's intent as expected.

When the utterances were produced with NS in two or more locations with a greater pitch change at a portion of the utterance other than the expected (Case 05), NS was perceived as expected more often when in initial position (34.8%, N = 8), followed by those in final position (32%, N = 8) and then by those in medial position (22.2%, N = 6). Note that utterances in Case 05 in medial position were perceived as initial 29.6% of the times and in final position 33.3% of the times.

Recall that the speakers tended to place NS on initial position together with NS in either medial or final positions, irrespective of the expected NS position. This tendency had different effects in the perception of NS among the listeners. One convergence strategy used by some listeners when there were pitch changes other than that in initial position was assuming that the pitch change in a position other than the initial would signal the NS while some listeners did not use same strategy. In other words, some listeners understood that initial pitch change was a trait of the speakers' speech rather than the speaker's intention of signaling important information, given that there were some other places being also stressed. Breakdowns or noise in communication happened with both types of listeners (using or not the convergence strategy), depending on the expected NS position.

When NS was expected in initial position and placed both in initial position and somewhere else (with a greater pitch), the listener who used the convergence strategy as described in the previous

paragraph (Case 05), misunderstood the speaker’s intent because the listener thought that it was not the speaker’s desire to place the NS in the beginning: It was just a trait of the speaker’s speech. Therefore, communication breakdowns happened. Figure 33 illustrates this occurrence.

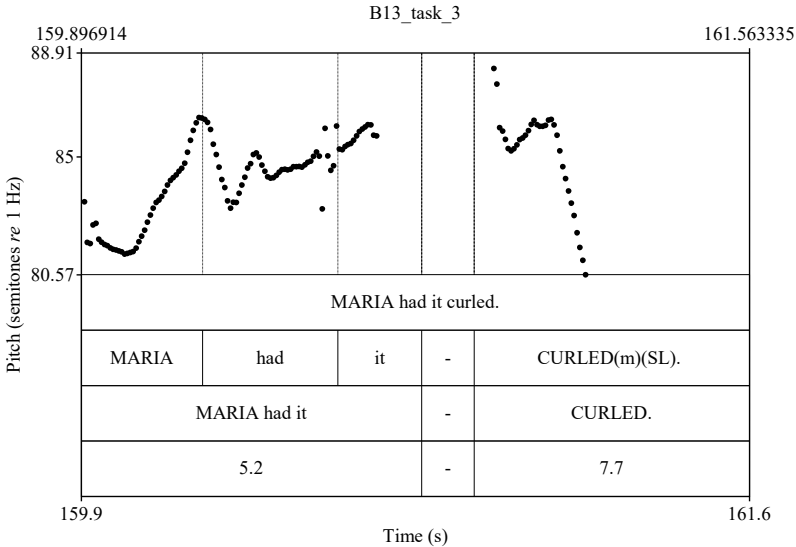


Figure 33: *MARIA had it curled* (Task 3, Item 09) by B13 – An example of a communication breakdown caused by the convergence strategy that pitch change in initial position was a trait of the speakers’ speech rather than the portion with important information.

In the utterance in Figure 33, NS should fall on *Maria* in order to imply that Maria rather than Ana had her hair curled. B13 had a wide pitch range and could produce some NS placements as expected. However, because of his inflections, many other words were salient. Additionally, some difficult words to pronounce seemed to receive more duration than expected with additional syllable lengthening, which might have the expected NS more difficult to be identified, given that both duration (Calhoun, 2007) and syllable lengthening (Shue et al., 2007) are also cues for NS. These traits of his speech are seen in Figure 33. Observe that he succeeded in placing NS on *Maria*. However, because of the hard word to pronounce ahead, *curled* (pronounced [ˈkɜːrlɪd] with syllable lengthening), he did a greater pitch change on the

word *curled*, which happened to have more duration. As F13 probably converged to the fact that B13 would make other words also salient, F13 might have considered the smaller pitch change on *Maria* as an unintended inflection and thus disregarded it as the information being corrected. As a result, he interpreted that B13's intent was to emphasize that Maria had her hair curled, not straightened.

When NS was expected in medial or final positions, some listeners (who did not develop the convergence strategy to accommodate to the fact that the pitch change in initial position was a trait of the speaker's speech) perceived the unexpected pitch change in initial position (greater than those in either medial or final positions) as signaling the most important piece of information, which led to real communication breakdowns. Figure 34 provides an illustration of this case.

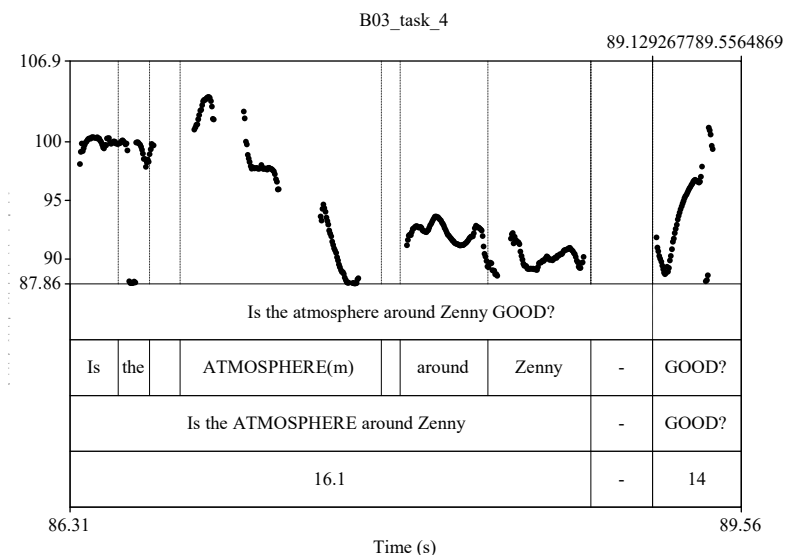


Figure 34: *Is the atmosphere around Zenny GOOD?* (Task 4, Item 05) by B03 – An example of a communication breakdown caused by the lack of convergence to the fact that pitch change in initial position was a trait of the speakers' speech rather than the portion with important information.

The NS in the utterance in Figure 34 should be assigned to the word *good* in order to imply that the speaker was in doubt about the

quality of the atmosphere when Zenny was around. Figure 34 shows that B03 did two pitch changes: one on *atmosphere* (pronounced ['æt̪məs, ferɪ]), with a 16.1-semitone-pitch change, and another one on *Zenny* (the expected location), with a 14-semitone-pitch change. There was a 298-millisecond pause before *good*. F03 interpreted that the initial pitch change had the important information and reported that pronunciation affected his judgement for this production. The listener had not still developed a convergence strategy in order to accommodate to the fact that the speaker would produce some pitch change in initial position and on difficult words, and thought that the final pitch change was only that of a question. Therefore, there was a communication breakdown here.

Table 44 considers these cases of unexpected productions, classified according to the position of NS placed by the speakers and the expected NS position, in relation to being perceived as expected or not.

Table 44

*Perception according to the unexpected productions of NS detailed according to position – general*

Production	Expected Perception		Unexpected Perception		Total	
	N	%	N	%	N	%
	(1) NS on a position other than the expected	3	20.0	12	80.0	15
(2) Nowhere	9	81.8	2	18.2	11	100.0
(3) 2 or more locations: greater on the expected	49	56.3	38	43.7	87	100.0
(4) 2 or more locations: no differences	36	31.9	77	68.1	113	100.0
(5) 2 or more locations: greater pitch change at an unexpected location	22	29.3	53	70.7	75	100.0

Table 44 displays the frequencies for the perception of unexpected productions that are classified into five categories regarding

position and the speaker's placement of NS, presented altogether, irrespective of being initial, medial or final: (01) placed at a position other than the expected (Table 42); (02) No NS along the utterance (plain pitch) (Table 43); (03) placed at two or more positions with the greatest pitch change on the expected position (Table 43); (04) placed at two or more positions with no differences in pitch changes (Table 43); and (05) placed at two or more positions with the greatest pitch change on a position other than the expected (Table 43). According to the figures, Cases 01, 04 and 05 (80%, 68.1%, and 70.7%, respectively) seemed to hinder communication to a greater extent than did cases 03 and 02 (43.7% and 18.2%, respectively). This is true mainly for initial, given that L1-BP possibly transferred this trait of the PB language into English – that of initial pitch change irrespective of the initial position having the important information (Carpes, 2014; Moraes et al., 2015). It means that these cases should be strongly avoided in communication between Brazilians of English and members of the international community when interacting in English. Learners should be taught to deemphasize unimportant information by destressing the corresponding words in the sentences, mainly those words placed in initial and medial positions.

This section discussed perception by the listeners in relation to the actual NS placements by the speakers. The section that follows addresses how these utterances were perceived when information was being elicited, corrected, and contrasted.

#### **4.2.3 RQ6 – NS Perception and speakers' intent interpretation according to type of information**

- RQ6: How is the speakers' intent interpretation affected by the type of information (being elicited, being corrected, being contrasted) in utterances?
- H6: Based on the findings in the pilot study and on the assumption stated in Hypothesis 2, Hypothesis 6 predicts that NS for signaling information being corrected and contrasted will yield more expected interpretations than NS for signaling information being elicited will.

Research Question 6 enquires about the effect of the types of information being conveyed by means of NS by L1-BP speakers of English have on the way NS is perceived and the speakers' intents are

interpreted by members of the international community. One of the findings of the pilot study was that unexpected interpretations were more frequent for unexpected productions of NS to provide information being elicited when compared to information being corrected. Therefore, based on this finding and on the assumption stated in Hypothesis 2 (Frota & Moraes, 2016; Klok et al, 2011), Hypothesis 6 predicted that NS placement to correct and contrast information would yield fewer misinterpretations of speakers' intent than NS to elicit information. This hypothesis is closely linked to Hypothesis 2, which predicted that the unexpected NS productions would be less frequent in utterances correcting and contrasting information. However, this hypothesis was not confirmed for production, given that contrasting information by means of NS showed to be statistically more difficult for Brazilians than correcting and eliciting information, as discussed in Section 4.1.2.

The present section now turns to the way these productions were perceived in order to understand if type of information played a role. Table 45 displays the perception results according to the type of information.

Table 45  
*Perception results according to type of information*

Status	Being Elicited		Being Corrected		Being Contrasted		Total	
	N	%	N	%	N	%	N	%
Expected	48	38.4	72	57.6	42	36.8	162	44.5
Unexpected	77	61.1	53	42.4	72	63.2	202	55.5
Total	125	100.0	125	100.0	114	100.0	364	100.0

The results displayed in Table 45 show that the expected perception of NS was more frequent for information being corrected (57.6%), followed by information being elicited (38.4%) and then by information being contrasted (36.8%). A chi-square test of independence was performed to examine if these differences were statistically significant. The same procedures for doing the test first described in Section 4.1.1 were used here. Tables 46 and 47 show the rearrangement of data and the partitioning of the contingency table in order to have two 2 x 2 tables instead of an only 2 x 3 table so the Chi-square test can be run.

Table 46

*Perception results rearranged for statistical testing (2 x 3 Contingency Table) – IV Type of information*

Information being	Expected Perception	Unexpected Perception
	N	N
Elicited	48	77
Corrected	72	53
Contrasted	42	72
Total	162	202

Table 47

*Perception results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Type of Information*

Information being	Exp.	Unexp.	Information being	Exp.	Unexp.
	N	N		N	N
Elicited	48	77	Elicited + corrected	120	130
Corrected	72	53	Contrasted	42	72
Total	120	130	Total	162	202

The results of the chi-square for these two partitions<sup>74</sup> in Table 47 are that, for the leftmost table,  $X^2(1, N = 364) = 9.231, p = .0024$  and, for the rightmost table,  $X^2(1, N = 364) = 2.873, p = .0901$ . These results indicate that the difference is statistically significant only for information being corrected.

The multiple comparisons with the  $p$  value recalculated with Bonferroni correction ( $p < .01$ ) confirmed the ones reported in the previous paragraph. Differences were statistically significant only for information being corrected if compared to information being elicited,  $X^2(1, N = 364) = 9.231, p = .0024$ , and being contrasted,  $X^2(2, N = 364) = 10.298, p = .0013$ , but they were not statistically significant between information being elicited and contrasted,  $X^2(2, N = 364) = 0.062, p = .8039$ . This means that for these listeners, perceiving NS on information being corrected and interpreting then the speakers' intents as expected was easier than on information being elicited and contrasted. Therefore, Hypothesis 6, which predicted that perceiving NS

<sup>74</sup> Recall that with the partitioning the df for each of the partitions is 1 + 1, which sums to the 2 df for the full table.

to correct and contrast information would be easier than to elicit information, was partially confirmed, given that only NS to correct information had statistically significant more expected perceptions and interpretations of speakers' intent.

The differences between information being contrasted and information being elicited did not reach statistical significance because they yielded very similar results. However, it is my impression that it was more an effect of structure rather than the communicative function itself, as discussed in Section 4.1.2. Information being contrasted shares with information being corrected the need for greater pitch change (Frota & Moraes, 2016; Klok et al, 2011). However, in this dataset, contrasting information showed to be a lot harder for speakers in the present study because the contrast was done by means of questions. Questions were by far the most difficult for the speakers to produce. Some of them had the knowledge that they should use a final rise in order to signal that it was a question and finished most of their questions with rising intonation. However, they would add some rising intonation somewhere else in the question followed by a low tone ( $N = 86$ ), as previously discussed in Section 4.1.2. In order for the listener to accommodate to these productions, more time was needed as the listener had to deal with many factors at the same time in order to develop appropriate convergence strategies (Giles et al., 1991): S/he had to understand if the rising intonation was there because of (1) the question, (2) the NS, (3) the difficult word being pronounced or because of (4) the presence of a difficult word to pronounce ahead.

Figure T20 (in Appendix T), reproduced here, illustrated a question with unexpected NS placement, combined with other cases, that was not perceived as expected by the listener. In that case, an unexpected production led to an unexpected interpretation and to a real communication breakdown. In order to understand when unexpected perceptions represented real communication breakdowns, Tables 48 and 49 were designed.



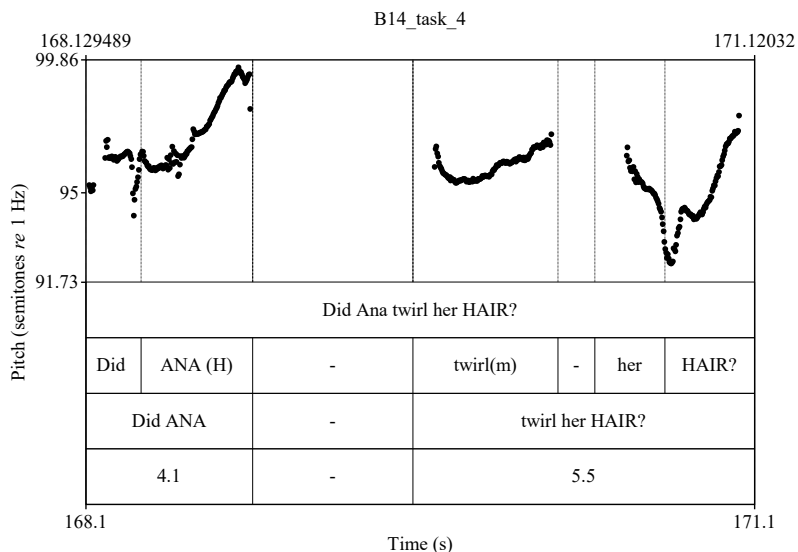


Figure T20: *Did Ana twirl her HAIR?* (Task 4, Item 26) by B14 – An example of an unexpected NS production combined with unexpected syllable lengthening/rising intonation, unexpected pauses, and unexpected pronunciation.

Table 48

*Unexpected perception results according to type of information – Repetition meaning reduced interpretability only vs repetition leading to real communication breakdowns*

Information being	Repeated production	
	N	%
Elicited		
Breakdowns	14	36.8
Reduced Interpretability	24	63.2
Total	38	100.0
Corrected		
Breakdowns	12	35.3
Reduced Interpretability	22	64.7
Total	34	100.0
Contrasted		
Breakdowns	16	38.1
Reduced Interpretability	26	61.9
Total	42	100.0

Table 49

*Unexpected perception results according to type of information – real communication breakdowns based on repetition Vs based on first production*

Information being	Breakdowns	
	N	%
Elicited		
Repeated	14	26.4
Not repeated	39	73.6
Total	53	100.0
Corrected		
Repeated	12	38.7
Not repeated	19	61.3
Total	31	100.0
Contrasted		
Repeated	16	34.8
Not repeated	30	65.2
Total	46	100.0

Table 48 displays the percentages for real communication breakdowns (generated when the speakers intent was not interpreted as expected, despite the production being heard once, twice or three times) and sites of reduced interpretability (that required the repetition of the utterance in order to reach the expected interpretation of the speakers' intent) according to the type of information (being elicited, corrected, and contrasted). Table 49 displays the percentage of times that real communication breakdowns happened even when negotiation of meaning happened (by means of repetition) and the percentage of times that they happened without negotiation of meaning (no repetition) for each type of information, meaning that neither the speaker nor the listener were aware that unsuccessful communication was taking place.

The figures in Table 48 show that the utterances that had sites of reduced interpretability (i.e., whose first production was not clear enough for the listener) were clear when heard for the second time more often in utterances with NS for correcting information (64.7%, N = 22), followed by those with NS for eliciting information (63.2%, N = 24) and then by those with NS for contrasting information (61.9%, N = 26), rendering communication successful in these cases, despite the noise caused. Note, however, that these figures are close to each other.

A little more of worry is the fact that more than 61% of the breakdowns in the three types of information represent the utterances

that were not repeated (73.6%, 65.2%, and 61.3% for information being elicited, contrasted, and corrected, respectively), as displayed in Table 49. Apart from statistical significance, this means that most real communication breakdowns happened when the listener was not aware of it, mainly in utterances with information being elicited, which goes in line with our initial prediction that NS for information being elicited would hinder communication to a greater extent as it does not require a pitch change range as large as correcting or contrasting information do.

It was not always the case that an unexpected production led to an unexpected perception and thus to a misinterpretation of the speakers' intent<sup>75</sup>, as can be seen in Table 50.

Table 50

*Perception results according to type of information and production status*

Information	Production	Perception	N	%
Being elicited	Unexpected	Unexpected	65	68.4
		Perceived as expected	30	31.6
		Total	95	100.0
	As expected	Unexpected	12	40.0
		Perceived as expected	18	60.0
		Total	30	100.0
Being corrected	Unexpected	Unexpected	48	47.1
		Perceived as expected	54	52.9
		Total	102	100.0
	As expected	Unexpected	5	21.7
		Perceived as expected	18	78.3
		Total	23	100.0
Being contrasted	Unexpected	Unexpected	69	66.3
		Perceived as expected	35	33.7
		Total	104	100.0
	As expected	Unexpected	3	30.0
		Perceived as expected	7	70.0
		Total	10	100.0

<sup>75</sup> For an illustration of an unexpected production leading to an unexpected perception, refer back to Figure 24.

The results displayed in Table 50 show that the utterances with the expected NS were interpreted as expected more frequently for information being corrected (78.3%, N = 18), followed by information being contrasted (70.0%, N = 7) and then by information being elicited (60.0%, N = 18). The utterances with the unexpected NS were interpreted as unexpected more often for information being elicited (68.4%, N = 65), followed by information being contrasted (66.3%, N = 69) and then by information being corrected (47.1%, N = 48). These differences were larger for information being corrected, which corroborates that NS in information being corrected, despite the utterance having noise, are easier to be understood than the other types of information.

This section answered Research Question 6, which sought to understand the role played by type of information on the perception of NS and the speakers' intent interpretation. It was predicted by Hypothesis 6 that NS for contrasting and correcting information would be more frequently perceived as expected than for eliciting information. This hypothesis was partially confirmed as only the perception of NS for correcting information reached statistical significance.

In Section 4.1.3, the analysis of production data showed that most utterances were produced with unexpected pronunciation, pauses and tones. The section that follows approaches their influence on the perception of NS and interpretation of speakers' intent by discussing the perception results for level of complexity of words in utterances: with non-challenging words, short-challenging words, or long-challenging words.

#### **4.2.4 RQ7 – NS Perception and speakers' intent interpretation according to level of complexity of words**

This section is divided into two parts. Section 4.2.4.1 presents the analysis of data in order to answer RQ7, which inquires about the influence of complex words on the perception of NS and the interpretation of speakers' intent. Section 4.2.4.2 draws comparisons between NS production and NS perception and interpretation of speakers' intent according to the complexity of words, and it addresses the effects of the combination of two or more of the pronunciation/prosody production aspects listed in Appendix T on perception.

#### 4.2.4.1 Answering RQ7

- RQ7: How is the speakers' intent interpretation affected by the complexity of words (non-challenging, short challenging, long challenging) in the utterances?
- H7: Based on the findings in the pilot study, complexity level of words in utterances (non-challenging, short challenging, and long challenging) will play a role in the way listeners interpret speakers' intent, as a result of the effect it will have in speakers' production, as predicted in Hypothesis 3. Utterances that contain challenging words, either short or long, will be more frequently misinterpreted than those that contain non-challenging words.

Research Question 7 enquires about the effect of the presence of complex words on the perception of NS and the speakers' intent interpretation by IL-English listeners. The pilot study showed that utterances with potentially difficult words to pronounce tended to yield unexpected productions and be misinterpreted by the listeners. The presence of words that are difficult for Brazilians to pronounce then became an independent variable investigated in the present study. It was hypothesized that utterances with either short- or long-challenging words would be misinterpreted more often than those with non-challenging words. Table 51 displays the results for perception according to the presence of these words in the utterances, irrespective of production status.

Table 51

*Perception results according to the complexity level of words in the utterances*

Status	Non-challenging		Short challenging		Long challenging		Total	
	N	%	N	%	N	%	N	%
Expected	62	51.2	51	41.8	49	40.5	162	44.5
Unexpected	59	48.8	71	58.2	72	59.5	202	55.5
Total	121	100.0	122	100.0	121	100.0	364	100.0

The results displayed in Table 51 show that the expected interpretation of speakers' intent was more frequent for utterances with non-challenging words (51.2%, N = 62) than for utterances with either

short- (41.8%, N = 51) or long- (40.5%, N = 49) challenging words, as predicted by Hypothesis 7.

In order to check if the difference is statistically significant, Tables 52 (one 2 x 3 table) and 53 (two 2 x 2 tables) were created in order to run the chi-square test for group independence (Larson-Hall, 2010).

Table 52

*Perception results rearranged for statistical testing (2 x 3 Contingency Table) – IV Complexity of words*

Words	Expected	Unexpected
	N	N
Non-challenging	62	59
Short-challenging	51	71
Long-challenging	49	72
Total	162	202

Table 53

*Perception results: Partitioning the 2 x 3 Table into 2 x 2 Tables – IV Complexity of words*

Words	Exp.	Unexp.	Words	Exp.	Unexp.
	N	N		N	N
Non-challenging	62	59	Non-challenging +	113	130
			Short-challenging		
Short-challenging	51	71	Long-challenging	49	72
			Total		
Total	113	130	Total	162	202

The results of the chi-square for these two partitions in Table 53 are that, for the leftmost table,  $X^2(1, N = 364) = 2.174, p = .1043$  and, for the rightmost table,  $X^2(1, N = 378) = 1.180, p = .2774$ . These results indicate that the difference is not statistically significant.

As in the previous sections, I opted for running multiple comparisons and thus I had the  $p$  value recalculated with Bonferroni correction ( $p < .01$ ). These multiple comparisons confirmed the results in the previous paragraph as none of the comparisons found a  $p$  value greater than .01: non-challenging  $V_s$  short-challenging,  $X^2(2, N = 364) = 2.174, p = .1043$ , and  $V_s$  long-challenging,  $X^2(2, N = 364) = 2.183, p = .0935$ ; and short-  $V_s$  long-challenging,  $X^2(2, N = 364) = 0.043, p = .8360$ . Therefore, Hypothesis 7 was not confirmed. This means that the

fact that utterances had the presence or absence of challenging words did not make NS perception difficult or easy for listeners to attain. However, as discussed in Section 4.1.3, this variable played a role for production, given that the complex words in terms of pronunciation triggered NS misallocations, productions of unexpected pauses and some other aspects that, undeniably, showed to distract some listeners as themselves reported. The next paragraphs provide some illustrations of these cases.

Figure T5, first used in Appendix T, provides an example in which a challenging word interfered with NS perception. For easy recall, it is repeated below.

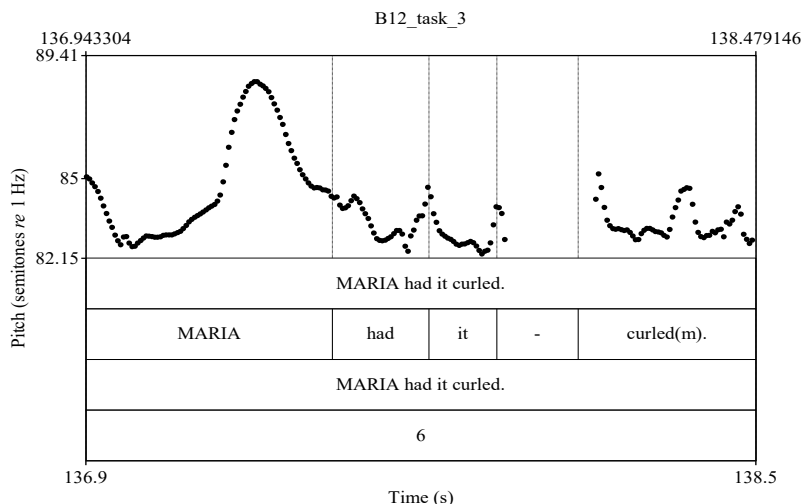


Figure T5: *No, MARIA had it curled* (Task 3, Item 09) by B12 – An example of an expected NS production with pronunciation issues in utterances with short-challenging words.

In the utterance in Figure T5, B12 successfully corrected that Maria rather than Ana had her hair curled, with a six-semitone-pitch change on *Maria*. There was a short pause right before *curled*, shorter than 200 ms and was not considered as a pause<sup>76</sup>. Although NS was

<sup>76</sup> As set in Section 3.4, in the present study only silence gaps longer than 200 ms were considered as pauses, given that these silences can be a result of the articulation for the production of certain sounds (e.g., /k/) and research has shown that such a short gap can not be perceived in fluent speech (e.g., Nootboom, 2007).

placed as expected, the word *curled* was pronounced [ˈkɜrlɪd], with the production of vowel epenthesis between /l/ and /d/. This unexpected way of pronouncing *curled* distracted the listener who asked B12 to repeat, signaling a site of reduced interpretability. Figure 35 shows the repetition of this production.

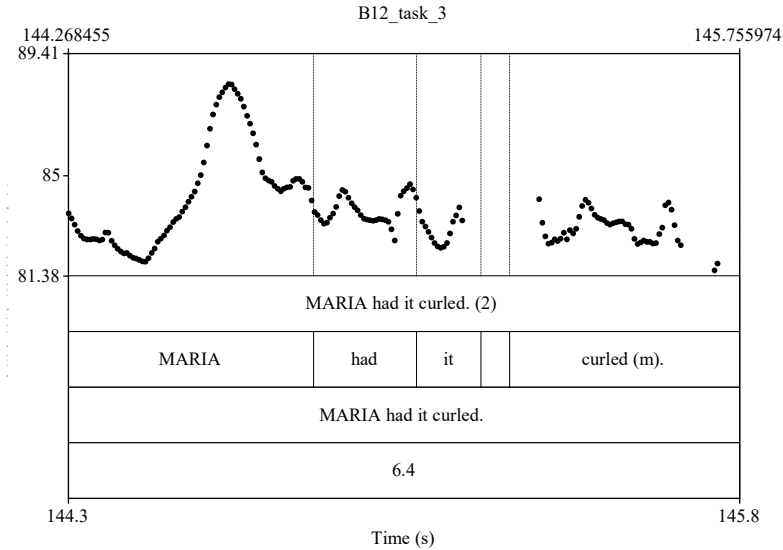


Figure 35: *MARIA had it curled* (Task 3, Item 09) repeated by B12.

In the repetition, her production was similar to the first with respect to pitch change (6.4 st) and pronunciation of *curled*. Nevertheless, F12 managed to (1) concentrate on the pitch change this time and (2) get the expected message. In this case, the unexpected pronunciation caused a site of reduced interpretability which was cleared out by repetition and thus did not lead to a real communication breakdown. In order to understand when unexpected perceptions were representative of sites of reduced interpretability or of real communication breakdowns, consider Tables 54 and 55.



Table 54

*Unexpected perception results according to complexity of words – Repetition meaning reduced interpretability only vs repetition leading to real communication breakdowns*

Complexity	Repeated production	
	N	%
Non-challenging		
Breakdowns	12	35.3
Reduced Interpretability	22	64.7
Total	34	100.0
Short-challenging		
Breakdowns	14	38.9
Reduced Interpretability	22	61.1
Total	36	100.0
Long-challenging		
Breakdowns	16	36.4
Reduced Interpretability	28	63.7
Total	44	100.0

Table 55

*Unexpected perception results according to complexity of words – real communication breakdowns based on repetition VS based on first production*

Complexity	Breakdowns	
	N	%
Non-challenging		
Repeated	12	32.4
Not repeated	25	67.6
Total	37	100.0
Short-challenging		
Repeated	14	28.6
Not repeated	35	71.4
Total	49	100.0
Long-challenging		
Repeated	16	36.4
Not repeated	28	63.7
Total	44	100.0

Table 55 shows that unexpected perception of NS when heard only once led to communication breakdowns more often in utterances

with short-challenging words (71.4%, N = 35), followed by utterances with non-challenging words (67.6%, N = 25), and finally by those with long-challenging words (63.7%, N = 28).

Table 54 shows that, when heard for the second or for the third times, unexpected perceptions were changed into expected more often for utterances with non-challenging words (64.7%, N = 22), followed by those with long-challenging words (63.7%, N = 28) and then by those with short-challenging words (61.1%, N = 22). Note, however, that these figures are close to each other. This means that the unexpected interpretations leading to communication breakdowns or to noise in communication were not affected by the level of complexity of words. It may be explained by the fact that even simple words were mispronounced by the speakers in the present study: *loves* (pronounced [lavʌs], [lɔvʌz], [l'ɔvɪz], [lav], [lavz], and [lʌvʌs]), *e-mail* (pronounced [i'meɪʊ], [i'meʊ], [i'meɪl], [i'meɪʊ], [i'meɪoʊ], and [i'meio]), *start* (pronounced [ɪs'tart], [ɪs'tarti], *like* ([laiki]), and *party* (pronounced [part], [pɑr], and [pɑtɪ]). These unexpected pronunciations affected listeners' perceptions in different ways. Some listeners made no mention to these pronunciations at all, while other listeners did (e.g., F01). Figures 36, 37, and 38 provide examples of the latter.

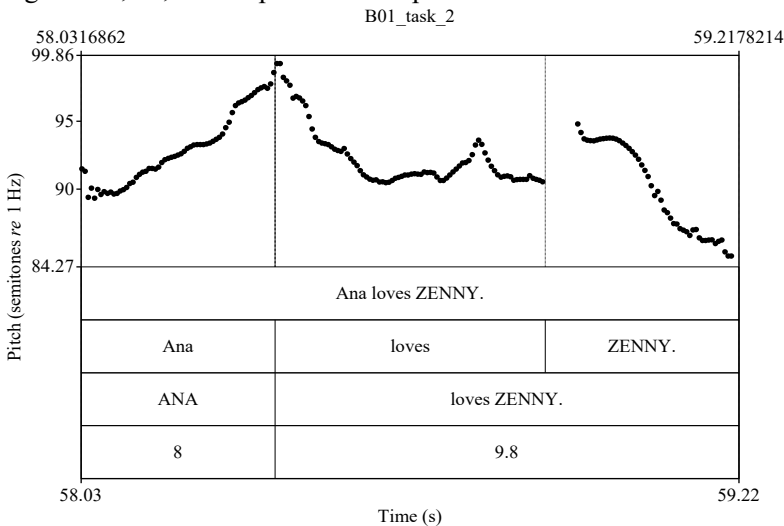


Figure 36: *Ana loves ZENNY* (Task 2, Item 05) by B01 – An example of unexpected pronunciations of non-challenging words affecting listeners' perception.

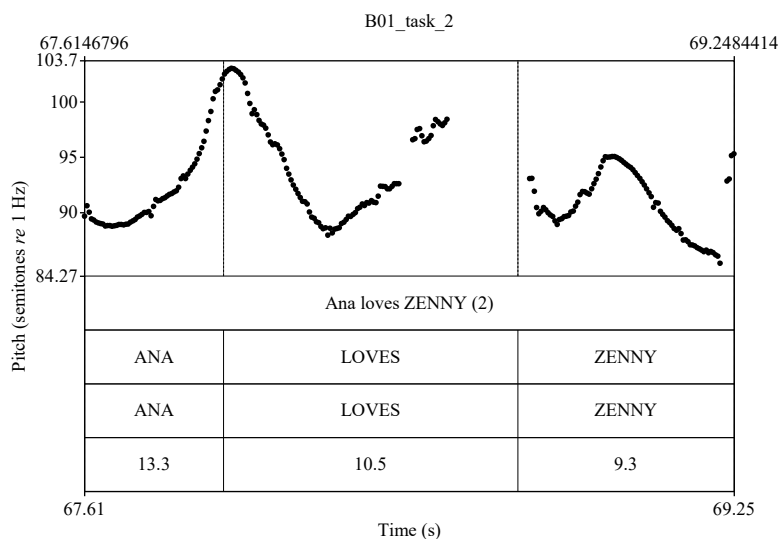


Figure 37: *Ana loves ZENNY* (Task 2, Item 05) repeated by B01.

The context question for the utterance in Figures 36 and 37 was *Who does Ana love?* Therefore, B01 had to nuclear stress the final portion of the utterance, namely *Zenny*. However, Figure 36 shows that she did two pitch changes: (1) on *Ana* (8 st with rising intonation) and *Zenny* (9.8 st). The word *loves* was pronounced [l'lovi:z]. Even though there was a greater pitch change on *Zenny*, F01 was attracted to the word *loves* and asked B01 to repeat.

In the repetition (Figure 37), B01 was a little louder and the Praat pitch settings had to be adjusted so that the pitch peaks could be displayed. Observe that this time B01 emphasized *Ana* (with a 13.3-semitone-pitch change and rising intonation), *loves* (pronounced [l'lovi:z], with a 10.5-semitone-pitch change and rising intonation), and *Zenny* (with a 9.3-semitone-pitch change and low boundary tone). F01 reported that she was distracted by the epenthesis in the word *loves*: “I would say *lovs* and not *lovis* and Brazilians say *lovis* and maybe therefore it gained my interest towards *loves* and therefore I thought that’s...the word with emphasis” (Interview with F01, 00:08:17 – 00:08:33). F01 ended up perceiving the emphasis on *loves* and interpreted that B01 was answering another question. Therefore, this was a real communication breakdown. Figure 38 illustrates the opposite.

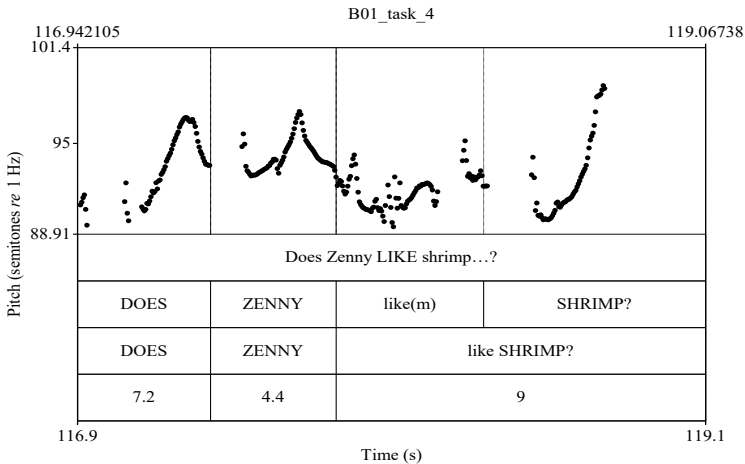


Figure 38: *Does Zenny LIKE shrimp?* (Task 4, Item 09) by B01 – An example of unexpected pronunciations of non-challenging words contributing to listeners' perception of NS.

In the utterance in Figure 38, B01 had to nuclear-stress *like* in order to imply that she was in doubt if Zenny liked shrimp or not. Observe that B01 pronounced *like* as [laɪki], a common case of epenthesis, but it attracted the listener's attention. It is interesting to note that the short-challenging word *shrimp* (which had the greatest pitch change and the longest duration) was pronounced as expected and that the word *like*, which should hold the NS, was the only destressed word in the whole question. The listener reported that the fact that *like* was mispronounced affected her perception. Luckily, *like* was expected to hold NS and in this case, mispronunciation helped the listener get the expected message.

This section answered Research Question 7 and found that the presence of challenging words in the utterances does not affect the way listeners' perceive NS and interpret speakers' intents. The reason for that may lay on the fact that even non-challenging words had unexpected pronunciations distracted some of the listeners participating in the present study.

Figures T5 and 35, discussed earlier in this section, illustrated an expected production in terms of NS placement with a pronunciation issue that caused an unexpected perception of NS. This perception turned out to be only a sight of reduced interpretability cleared out after

the utterance was heard for the second time, and it thus yielded a successful interpretation of the speakers' intent. Nonetheless, it was not always the case. The next section compares production and perception related to complexity of words in the utterances.

#### **4.2.4.2 Perception results according to actual NS productions and complexity of words**

In order to have a clearer picture of the perception according to expected and unexpected productions in utterances with non-, short-, and long-challenging words, consider Table 56. It displays the results for the way the NS was perceived whether or not placed as expected, taking into consideration the utterances with the three types of words: non-challenging, short challenging, and long challenging. Expected productions were more often perceived as expected in utterances with short-challenging words (80.0%, N = 12), followed by utterances with non-challenging words (72.7, N = 24), and then by those with long-challenging words (46.7%, N = 7). Note that the difference in terms of production, presented in Section 4.1.3, showed that expected productions were more frequent in utterances with non-challenging words. However, as mentioned in Section 4.2.4.1, even the non-challenging words, which were mostly short words, were produced in unexpected ways, for example, with the production of epenthesis (e.g., *start*) and vowel change (e.g., *loves*). This may in part explain why the perception results of expected productions yielded similar figures. The difference was clearly larger for long-challenging words, mainly *hierarchy*, whose pronunciation showed to be a challenge to all speakers in the present study. This definitely points to the importance of dealing with the pronunciation of long words during learning English so they do not become obstacles when speakers communicate in English.

Table 56

*A comparison between production and perception – complexity of words*

Utterances with	Production	Perception	N	%
Non-challenging	Unexpected	Unexpected	50	56.8
		Perceived as expected	38	43.2
		Total	88	100.0
	As expected	Unexpected	9	27.3
		Perceived as expected	24	72.7
		Total	33	100.0
Short challenging	Unexpected	Unexpected	68	63.6
		Perceived as expected	39	36.4
		Total	107	100.0
	As expected	Unexpected	3	20.0
		Perceived as expected	12	80.0
		Total	15	100.0
Long challenging	Unexpected	Unexpected	64	60.4
		Perceived as expected	42	39.6
		Total	106	100.0
	As expected	Unexpected	8	53.3
		Perceived as expected	7	46.7
		Total	15	100.0

Unexpected productions led to unexpected perceptions almost to the same extent for utterances with non-challenging (56.8%, N = 50), long-challenging (60.4%, N = 64), and short-challenging words (63.6%, N = 68). In order to understand what sort of pronunciation issues were involved in the ones perceived as expected and the ones perceived as unexpected, refer to Appendix U.

Table 57 displays the total counts of expected NS placements (Column *Prod.*), and perceptions (Column *Percep.*) for each utterance in the tasks (Column *Utterances*). It also displays the pronunciation mean for each word as rated by the speakers (Column *M.Pron.*). The

utterances are grouped according to the level of complexity of words and the type of information is provided in the Column *Info.Type* (E. = elicited; Cor. = corrected; Con. = contrasted).

Table 57

*Total counts of expected productions and expected perceptions for each utterance*

	Utterances	Info. type	Prod.	Per cep.	M.P ron.
Non-challenging	ZENNY loves you.	E.	5	9	8.5
	Zenny LOVES you.	E.	5	5	8.5
	Ana loves ZENNY.	E.	3	4	8.5
	No, ANA sent the e-email.	Cor.	6	11	*
	No, Ana LOVES dogs.	Cor.	9	12	8.5
	No, Zenny loves DOGS.	Cor.	0	7	8.5
	Do MEN lie a lot?	Con.	0	8	8.3
	Does the meeting START at ten?	Con.	0	3	8.6
Does the party start at TEN?	Con.	5	3	8.6	
Short-challenging	ZENNY sent you the <b>pearls</b> .	E.	4	7	5.3
	The boss LOVES your <b>world</b> .	E.	3	4	5.5
	The T-shirt <b>shrank</b> A LOT.	E.	0	5	5.7
	No, MARIA had it <b>curled</b> .	Cor.	3	7	3.9
	No, <b>pearls</b> MELT in <b>vinegar</b> .	Cor.	0	5	5.0
	No, the <b>rhythm suits</b> ZENNY.	Cor.	1	9	5.8
	Is ZENNY afraid of <b>heights</b> ?	Con.	1	5	5.3
	Does Zenny LIKE <b>shrimp</b> ?	Con.	1	6	7.1
Did Ana <b>twirl</b> HER HAIR?	Con.	2	3	2.4	
Long- challenging	ZENNY <b>traveled</b> to <b>Pennsylvania</b> .	E.	6	5	7.0
	Zenny DISLIKES <b>vinegar</b> .	E.	3	3	4.7
	The <b>government</b> talked to ZENNY.	E.	1	6	6.1
	No, RED is an <b>appetite</b> stimulator.	Cor.	3	9	6.3
	No, yellow STIMULATES <b>optimism</b> .	Cor.	1	5	5.9
	No, <b>unfortunately</b> Zenny AGREED.	Cor.	0	7	8.4
	Is ZENNY leaving <b>immediately</b> ?	Con.	0	4	6.3
	Does this sequence REPRESENT the <b>hierarchy</b> ?	Con.	1	4	3.5
Is the <b>atmosphere</b> around Zenny GOOD?	Con.	0	6	6.8	

\* Not rated.

If the production column in Table 57 is considered, one will see that 8 out of the 27 utterances in the tasks did not have NS placed as expected: *No, Zenny loves DOGS*; *Do MEN lie a lot?*; *Does the meeting START at ten?*; *The T-shirt shrank A LOT*; *No, peals MELT in vinegar*; *No, unfortunately Zenny AGREED*; *Is ZENNY leaving immediately?*; and *Is the atmosphere around Zenny GOOD?* It is interesting to observe that these are utterances with non-challenging words (N = 3) as well as utterances with challenging words, either short (N = 2) or long (N = 3). Additionally, note that the pronunciation rating by the speakers about how intelligible they thought they would sound when pronouncing these words was quite distinct, ranging from 5.0 to 8.6.

As seen in Section 4.1.3, the words considered the least intelligible were *curled* (M = 3.9), *hierarchy* (M = 3.5), *twirl* (M = 2.4), and *vinegar* (M = 4.7). Utterances containing these words also had low total counts for production: 3, 1, 2, and 3, respectively. It was also true for perception: 7, 4, 3, and 3, respectively. Indeed, as can be seen in Table 57, NS productions were mostly unexpected, and the reasons why this was so can be manifold: a simple phonological process such as vowel epenthesis might have influenced some while the interplay of many processes might have influenced others, as discussed along this chapter.

Recall that, in Section 4.1.3, it was said that the difficulties posed by the presence of challenging words caused speakers to produce unexpected placements, unexpected pronunciations, unexpected pauses, unexpected rising intonation, unexpected syllable lengthening and repetition. In order to understand if the combination of these aspects impaired communication to a greater extent than the presence of a sole aspect, Table 58 was devised. It summarizes the results for the perceptions according to the number of processes involved in the production of the utterances. It puts together in the first row the results for perceptions of unexpected productions that had no pronunciation problems or had only one problem during its uttering, be it related to pronunciation of a word alone, or pause alone or syllable lengthening alone. The second row displays the sum of the results for perceptions of unexpected productions that included two or more aspects.



Table 58  
*Number of aspects in unexpected productions and perception: a comparison*

Status	Exp.	Red.	Breakdow	Total	Total
	Percep.	Interpret.	ns	Unexp.	
	N	N	N	N	N
	(%)	(%)	(%)	(%)	(%)
None or only one aspect	61 (51.3)	24 (42.9)	50 (39.7)	74 (40.7)	135 (44.9)
2 or more aspects	58 (48.7)	32 (57.1)	76 (60.3)	108 (59.3)	166 (55.1)
Total	119 (100)	56 (100)	126 (100)	182 (100)	301 (100)

The figures in Table 58 shows that 60.3% (N = 76) of the communication breakdowns resulted from utterances produced with two or more processes and that reduced interpretability was also more frequently a result of utterances with two or more processes (57.1%, N = 32). These results indicate that the combination of two or more aspects make the speech of these speakers more difficult to process and their intent more distant of being grabbed. The differences for the expected perceptions were not far from each other, but the expected perceptions were more frequent in utterances produced with none or only one of the aspects described (51.3%, N = 61).

These figures may indicate that some listeners may have developed convergence strategies to accommodate to the speakers' speech as follows:

- 1) When NS on a word was perceived and the word was mispronounced, listeners suspected that the word would probably not be the intended place for NS and that the speakers were only having difficulties to utter the hard sentence. As a result, they would ask the speakers to repeat in order to try to notice NS somewhere else in the utterance. In some cases, the NS was noticed and the intent was successfully interpreted, while, in other cases, the production remained the same and the intent was totally misinterpreted. The listener assumed that the speaker sometimes emphasized a word because it was difficult, but not because it was important.

- 2) When the utterance had a difficult word pronounced as expected and the word had some pitch change probably because the speakers were trying to be careful when pronouncing it, suspicion about the difficult word holding the NS was not always raised and most cases led to communication breakdowns and only some to reduced interpretability. In cases like this (Aspect 10, Appendix U), most listeners assumed that if the speaker emphasized a word and did not mispronounce it, the difficult word was probably the one that should have the NS, because it was important, not because it was difficult.

Overall, these aspects that emerged from production when speakers were attempting to place NS as expected should be avoided for some reasons.

First, unexpected pronunciation affects not only the perception of the segment but also of suprasegmentals, as depending on the phonological process involved, such as the case of vowel epenthesis, a new syllable can be formed, causing word stress shifting, which may have an impact on NS if the word is nuclear-stressed. If the unexpected pronunciation is in a destressed word, problems may also arise, as discussed in Appendix T. It was explained that many listeners reported that they could notice the NS on a specific word, but they would not be able to understand the remainder of the utterance. One example of this case is the production of *Is Zenny afraid of heights?* (Task 4, Item 03) by B07. The word *heights* was mispronounced. F07 had understood *AIDS*, thought that the combination was strange – *Is Zenny afraid of AIDS or dogs?* – and suspected that it was not the correct pairing and asked B07 to repeat:

I also asked because it was like... a strange... comb... *is she afraid of AIDS or dogs?* Like I, I expected more like *Is she afraid of CATS or dogs?* Like...I, because it's like ... hum... it was so out of ... like out of category (...) [researcher explains B07 meant *heights* and how the production should have been]. I was thinking that she said AIDS, you know, the, the disease? Like *AIDS*... it was so strange... (Interview with F07, 00:19:25 – 00:20:39).

It is important to note that B07 nuclear stressed *Zenny*, which accorded with the discursive context. However, because of the unexpected pronunciation of *heights*, which led to the recognition of the word as *AIDS*, F07 was unable to figure out if the NS was really on *Zenny*. After B07 repeated the utterance, F07 filtered out the pronunciation and the awkwardness of the utterance and got it correct. This specific example made me wonder about Jenkins's (2000) example of a communication breakdown in which an IL-English speaker nuclear-stressed the wrong constituent in a question, which was also mispronounced: "Do you have a blue *vun*?", *vun* meaning *one*. She argues that if *vun* had been destressed, no misunderstandings would have taken place. However, as I see it, based on some of the interactions of the present study, this is not so straightforward, as pronunciation of constituents other than the one where NS falls seemed to be also important to communication.

Second, having in mind that pauses and syllable lengthening (and rising intonation) are also cues for thought group limits (Heusinger, 2007), the production of unexpected pauses and unexpected syllable lengthening made it difficult for listeners' to identify the NS as more than one NS was produced in many cases due to the presence of many thought groups.

Lastly, syllable lengthening alone is also a cue for NS given that the nuclear-stressed constituent has longer duration (Keating et al., 2003; Moraes et al., 2015; Shue et al., 2007). Therefore, unexpected syllable lengthening on a constituent that was not meant to be nuclear stressed confused the listeners who relied on duration in order to disentangle difficult productions.

#### 4.2.5 Summary of perception results

This section presented and discussed the perception results according to the location of NS, the type of information, and level of complexity of words in terms of pronunciation of the utterances in Tasks 2, 3, and 4 (presented in Section 3.3), where information had to be elicited, corrected, and contrasted, respectively.

Perception results showed that listeners did not perceive NS as expected most of the times (Section 4.2) and these unexpected perceptions resulted in communication breakdowns more frequently (64.4%) than in only reduced interpretability (35.6%), that is, noise in

communication that was lowered down after negotiation of meaning by means of repetition and led to successful communication.

Hypothesis 4, which predicted that unexpected NS placements would hinder the perception of NS and the interpretation of speakers' intents, was confirmed. However, it was discussed that unexpected productions also yielded a considerable number of expected perceptions (Section 4.2.1). This finding goes in line with previous studies in the area (Atechi, 2004; Tiffen, 1974) that found that unexpected productions led to unexpected perception of NS and speakers' intents, while it also problematizes this straightforward correspondence.

Hypothesis 5, which predicted that utterances with NS in final position would yield unexpected perceptions more often than would those in either initial or medial positions, was not confirmed, given that expected or unexpected perceptions were not statistically different for the three utterance positions (Section 4.2.2.1). A comparison of the perception of the actual assignment of NS (Section 4.2.2.2) showed that expected NS placement in initial position was a little easier to perceive than in final and medial positions<sup>77</sup>, but differences were close to each other. However, the production cases that emerged from data analysis seemed to interfere with the perception of NS and speakers' interpretation, mainly for NS placed on a location other than the expected (Case 01), for NS placed in more than one location, with no distinction between pitch changes (Case 04) or with a greater pitch change on an unexpected location (Case 5). These unexpected pitch changes were more frequent in initial position and it might be the case speakers transferred that pitch movement from their L1 (Carpes, 2014; Moraes et al., 2015).

Hypothesis 6 predicted that NS for correcting and contrasting information would yield more expected perceptions. It was partially confirmed, given that the differences only for information being corrected reached statistical significance (Section 4.2.3). Contrasting information was done by means of questions and that might have interfered with the results as English questions are very challenging for L1-BP speakers of English to produce (Passarella-Reis, 2014). This finding for NS in utterances for correcting information corroborates

---

<sup>77</sup> This finding may be related to the fact the L1-BP speakers might have transferred to English their tendency to compress pitch of unfocused material when important information is on initial position in PB utterances (Carpes, 2014).

studies that view pitch change for correcting information as wider than NS in those for eliciting information (Frota & Moraes, 2016).

Hypothesis 7, which predicted that NS in utterances with challenging words would be more frequently misperceived, was not confirmed. Although there were differences, they did not reach statistical significance (Section 4.2.4.1). This finding may be attributed to the fact that even non-challenging words had minor unexpected pronunciations, which surprisingly interfered with the speakers' perceptions. A comparison of their actual production of the utterances and the way they were perceived (Section 4.2.4.2) corroborated the reason mentioned in the previous period and showed that unexpected NS assignments in utterances with long-challenging words led to more unexpected perceptions. The aspects that emerged in the analysis of production data in relation to the complexity of words showed that short- and long-challenging words caused the combination of more aspects in the productions, which, in turn, showed to affect perception leading to unexpected interpretations a little more often than when only one aspect was involved. The differences were not great for perception because listeners developed strategies in order to converge to speakers' speech, which helped in some utterances and hindered communication in others. This was so because some production aspects (e.g., unexpected pauses and syllable lengthening/duration) are also important cues to NS (Keating et al., 2003) or thought group (Heusinger, 2007) identification.

#### 4.3 CONCLUSION OF THE CHAPTER

First, the present chapter presented the production results (Section 4.1). The main findings for production of NS were the following: L1-BP speakers of English had difficulties to place NS as expected, especially in (1) utterance final position, given that they did not destress unfocused material in initial and medial position (Section 4.1.1), (2) questions contrasting information, probably an effect of language transfer (Section 4.1.2), and (3) utterances containing complex words in terms of pronunciation (Section 4.1.3).

Second, the chapter presented the perception results (Section 4.2). One of the findings for the perception of NS was that IL-English listeners had difficulties to perceive the NS and interpret speakers'

intent as expected, especially when NS was not placed as expected, but, due to the development of convergence strategies, they also managed to succeed in their interpretations even when NS was not expectedly placed (Section 4.2.1). Another finding was that expected NS position does not interfere with the perception of NS, but actual NS position does: NS in initial positions were a little easier to perceive (4.2.2). Finally, it was also found that the presence of complex words in utterances does not play a role for perception, but the actual pronunciation of the words does (4.2.3).

Smith and Nelson's (1985) stand for intelligibility says a lot about the way BP-L1 speakers of English were understood in their intents by the IL-English listeners: Intelligibility is interaction-dependent, and influenced (a) by experience with either the interlocutor's accent or the interlocutor's language (belonging to a given speaking community), and (b) by the listener's expectations on (positive attitude and willingness to understand the interlocutor's speech).

Language is dynamic and so is understanding and producing language. As this study advanced in the analysis of the dataset, it was noticed that each pair of speaker/listener had a different interaction. Accommodation to the speech of speakers happened in different manners for different listeners. For instance, for some listeners, a word *preceded* by a pause was seen as the nuclear-stressed word, while for others the word *followed* by a pause was seen as the nuclear-stressed word. On the one side of the coin, there are the listeners, unique human beings with personal experiences that influence the way each of them processes and perceives language. On the other, there are the speakers, who carry with the speakers' self their language insecurities, personality traits, voice qualities, anxiety or lack of it, which influence their performance and measuring this influence is this far not possible. This study now understands that looking into each interaction separately would be of great value. In all, unexpected ways of producing language represent obstacles, big or small, which may lead to annoying and stressful interactions up to real breakdowns in communication.

## 5 FINAL REMARKS

The present study aimed at contributing to the body of research that is dedicated to (1) investigating the effect of unexpected prosody in international communication and (2) informing pedagogy on the most important pronunciation features for the intelligibility of IL-English speakers' speech. It investigated the use of NS by 14 L1-BP speakers of English as the only cue to signal the most important information in an utterance. It also investigated the perception of NS placed by L1-BP speakers of English by 14 members of the international community (IL-English listeners) and the resulting interpretation of these speakers' intent.

The research questions encompassed three independent variables: location of NS (initial, medial, and final), type of information according to the function played by the NS (for eliciting, correcting, and contrasting information), and complexity of words (non-challenging, short-challenging, and long-challenging).

The production dataset was composed of 501 utterances (378 first-time productions and 123 repetitions) produced by the speakers during interactions with the listeners in three pairwork tasks, and the video recordings of the tasks. The perception dataset was composed of 378 listeners' answers in the tasks sheets and the listeners recorded comments in the follow-up interviews.

The analysis of data showed that L1-BP speakers had difficulties to convey important information by means of NS especially if the information was located in the rightmost end of the utterance, as investigated by Research Question 1. This may be a result of language transfer given that in their L1, when important information is in final utterance position, other pitch movements are also found in the beginning and/or middle of the utterances (Carpes, 2014). The perception results showed that only the fact that NS is expected in initial, medial or final position does not affect its perception. However, the perception results that took into consideration the actual NS locations assigned by the speakers showed that NS in initial position were easier to perceive. This finding corroborates research saying that NS pitch change in non-final positions do not require a great pitch change range in order to be perceived if compared to NS in final utterance positions (Keating et al., 2003).

Regarding NS placements in order to convey the three types of information investigated (RQ2), it was found that for these speakers contrasting information by means of questions was more difficult if compared to eliciting and correcting information in statements. This finding was attributed to the fact that for L1-BP speakers, asking questions in English is a challenge in intonational terms (Passarella-Reis, 2014). Given that English questions finish with a final pitch rising movement, speakers tended to do that movement but always departing from the last constituent of the question, irrespective of it holding or not the most important information, which led to unexpected productions. With respect to the perception of NS and the interpretation of speakers' intent, NS used to signal information being corrected showed more expected perceptions and interpretations. This finding corroborates research saying that pitch change range is wider and thus easier to perceive in NS for correcting information than in NS for eliciting information (Frota & Moraes, 2016).

As for the effect of the presence of complex words in utterances, it was found that both short- and long-challenging words hindered the expected NS placement, given that their presence caused more anxiety and more care when speakers were about to say or while they were saying the utterances, which made arise other pronunciation aspects such as unexpected pauses, unexpected syllable lengthening/rising intonation, breaking speech in more thought groups and making arise more NSs. In terms of perception, the simple fact that utterances had complex words did not impact the way NS was perceived and the speakers' intents were interpreted. However, a comparison between production and perception according to the presence or absence of these words showed that utterances with long-challenging words were misperceived more often.

Unexpected NS placement caused noise to communication or even led to communication breakdowns (RQ4). However, it was found that, although the unexpected assignments of NS more often led to the unexpected interpretation of speakers' intents, they also led to the expected interpretation by means of negotiation of meaning (resorting to repetition) and of the development of convergence strategies by the listeners (Giles et al., 1991).

Speakers developed different production strategies to deal with difficult words: using plain tone (F14), pausing before the word, using syllable lengthening or rising intonation on the word that preceded the



hard word on the hard word itself, using or not extra cues such as eyebrow movement or nodding together with the delivery of the NS word. The convergence strategies developed by the listeners to deal with these production strategies worked for some interactions and differed among listeners participating in the study. The production strategy used by B14 of distressing hard words to pronounce was probably the most successful one in order to help the listener find the NS in her productions. Note, however, that it does not mean that the listener would be able to understand what she was saying in the pitch-compressed part of the utterance, which itself would be cause of misunderstandings.

It is strikingly important to highlight that, although position, type of information and the presence of complex words were variables investigated in the present study, the reasoning guiding this study is that language is a dynamic system and so is communicating through language, which has an interplay of many aspects intertwined and combined with each other. Consequently, it is not auspicious to affirm that one factor alone would be responsible for breakdown in communications.

## 5.1 PEDAGOGICAL IMPLICATIONS

The results discussed for each independent variable for production offer fractal information that gives away what the general results in Table 13 (Section 4.1) showed: Placing NS for these Brazilians was a difficult and challenging task. During data collection, when receiving the explanation about the tasks, most of them asked the researcher how they were supposed to emphasize the information, a piece of evidence that they were not aware of how this phenomenon would take place. Based on Jenkins's (2000) assertion about the teachability of NS, the fact that Brazilians are not taught how to place NS in English utterances, and that they have limited awareness of NS (Kivistö-de-Souza, 2017) the insertion of explicit instruction of NS in regular English classes is advised. This instruction would raise Brazilians' awareness of the implications of unexpected/unintended relevant pitch changes to their speech when communicating in English.

Unexpected pauses emerged from speakers' speech. These pauses represented a threat to communication in terms of prosody processing, as they are cues for thought group limits (Heusinger, 2007) and they may break speech in nonsense units. Breaking speech into meaningful

units by avoiding unnecessary pauses should be also an aim in English classes.

Three aspects that emerged from data analysis are worth noting. The first is that common processes such as vowel epenthesis surprisingly affected NS perception and the speakers' intent interpretation. One possible explanation to this is that epenthesis is considered a reinforcement resource (Migliorini & Massini-Cagliari, 2010) and this being so the listener felt that the epenthesized word would hold the NS. The second aspect refers to syllable lengthening. This phenomenon would be expected on the NS word as more prominent words usually also have more duration (Calhoun, 2007; Shue et al., 2007). However, in the present dataset, it was also a resource used by speakers to gain more time or it was used in words they were careful when pronouncing. The third aspect is that in *ZENNY loves you*, *Zenny LOVES you*, and *ZENNY sent you the pearls*, it happened that *you* was perceived as holding the NS even when *you* had the smallest pitch change in the utterance. This may be explained by the fact that function words are usually destressed and when they are the constituent holding the important information (focused word), NS does not take a great pitch change in order for the change to be perceived (Calhoun, 2007), and this misguided listeners' judgements in some cases. These three aspects together are related to the importance of destressing unimportant material and points to the significance of teaching L1-BP speakers of English to destress unfocused information. Having in mind previous studies have shown that destressing is teachable and acquired before the English NS rule (Nava & Zubizarreta, 2010), L1-BP speakers would greatly benefit if destressing were explicitly taught in English classes.

Jenkins (2000) does not include word stress as an important feature in the Lingua Franca Core. Note that long-challenging words made arise more unexpected perceptions of NS than did short-challenging words. The difference between these variables is number of syllables and word stress (long words may have the fourth syllable stressed, a rare stress pattern in PB). Dealing with stress in long words is one of the challenges faced by L1-BP speakers of English (Brawerman, 2006). This may lead us to infer that word stress is also important as it has some impact on how speech is produced and perceived. This definitely points to the importance of dealing with the pronunciation of long words during learning English so they do not become obstacles when speakers communicate in English.

## 5.2 LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

I would like to mention that the method used in the present study triggered the full engagement of both speakers and listeners into the tasks designed for data collection. This engagement resembled that of people involved in real communication, where communicating information is the ultimate goal. As a result, although it used reading, the speaker was more concerned with the message than with the decoding of the utterances. Nevertheless, research in the field of prosody has not used this kind of methodology for data collection. Future prosody studies that need to have some variables under control would benefit from this kind of method.

One of the limitations of the present study is related to the use of questions. Having information being contrasted in questions was not the best choice to check NS placement in order to contrast information. Even though the first clause of disjunctive questions in both English and BP languages share the same boundary tone (H%), asking questions for Brazilians was still difficult, as shown in previous studies (e.g., Passarella-Reis, 2014). L1-BP speakers seem to transfer other prosodic traces of BP to English, such as the lack of pitch compression in the beginning of the question when NS is in final position and the non-maintenance of the high tone until the end of the question when NS is on non-final position. Future studies should investigate contrasts being made in statements to neutralize the effect of intonational patterns and so that results be comparable in the three levels of the independent variable (type of information).

This study gathered the reasons reported by the listeners for interpreting or misinterpreting the speakers' intent. Future studies could investigate if the reasons given by the listeners match the actual speakers' productions, in order to disentangle this interesting relationship between perception and production.

In order for a pitch change to be considered perceivable in speech, the present study took the stand that a three-semitone change in pitch was necessary (t'Hart, 1981). Therefore, only pitch changes greater than that were classified as locations having NS. However, in the dataset, pitch changes smaller than that showed to be perceived by some listeners. Future studies should work with the possibility of following literature that sees a one-and-a-half-pitch change as perceivable in terms

of prominence (e.g., Rietveld & Gussenhoven, 1985 as cited in Nootboom, 1997). Analyzing prosodic data by using a lower pitch change range parameter may yield strikingly different results in terms of both production and perception of NS.

In the present study only silence gaps longer than 200 ms were considered as pauses, given that these silences can be a result of the articulation for the production of certain sounds (e.g., /k/) and research has shown that such a short gap cannot be perceived in fluent speech (e.g., Nootboom, 2007). However, some listeners in the present study perceived some short pauses in contexts followed by other sounds rather than the voiceless stops. Future studies should consider the possibility of such shorter gaps not followed by voiceless stops be taken as pauses.

NS analysis relied on the physical correlate of pitch, that is, the fundamental frequency. Even though an impressionist account of duration was carried out, the present study did not use it as a consistent parameter. Future studies investigating NS should benefit from duration as well as from loudness and pitch peak alignment to look into NS production and perception.

Some assertions about language transfer were made in the present study. However, they should be further investigated as these speakers' speech in their L1 was not investigated in order to make direct comparisons with their productions in English. There is a dearth of studies in this area in Brazil and research in this field should be fruitful.

It was reported that there were cases in which the listener could perceive the emphasis but the listener was unable to understand the remainder of the utterance. This just gave us food for thought. Recall that Jenkins (2000) gives an example of an IL-English speaker interacting with other members of the international community and asking for a blue pen, by saying "Do you have a blue vun?", *vun* meaning *one*. Jenkins argues that this speaker's interlocutors could not make sense of what was being said because *vun* held the NS. If the speaker had placed the NS on blue (in order to make a contrast with any other color), she argues, the other interactants would have understood that he wanted a blue pen. However, my point here is that it might not be the case. If the listeners in my study did not have the options, would they have been able to grab the speaker's intent? The case discussed and illustrated in Figure T6 (Appendix T) points to a no. The listener understood only the word being corrected (*red*), so he got that the speaker's intent was to correct the color, but he was not able to

understand the message at all. Who can guarantee that those interactants in Jenkins's study would have understood right away what was being said if *blue* were nuclear-stressed? It is possible that as they did not have any options to check, they would try to understand the whole question and getting the meaning of *vun* was important. Future studies could disentangle this issue by having utterances being heard by two groups: with and without the options.

Future studies involving the interpretation of speakers' intent should be case studies. Putting together different listeners does not seem to enable a clear understanding of what goes on during the process of accommodating to one's speech.

The present study did not control for L1 background as it wanted to investigate how Brazilians' intent would be interpreted by the international community irrespective of listeners' nationalities. However, I am aware that the L1 of each listener may have also affected the way that they perceived speech and that they developed convergence strategies in the present study. Future studies grouping participants according to their L1 background would raise an interesting discussion on NS perception mainly in cases in which NS is not placed as expected.

Level of proficiency was controlled in the present study due to the nature of the tasks involved in data collection and it was not the objective to draw comparisons between different levels of proficiency (e.g., pre-intermediate speakers *Vs* advanced speakers) regarding the use of NS. However, participants did differ in their levels despite the fact that they were all above pre-intermediate level. During the analysis of data, I noticed that proficiency might be an interesting variable to consider. Less proficient speakers and listeners tended to have different behaviors if compared to more proficient speakers and listeners. Less proficient speakers exhibited more frustration towards their own speech (e.g., B01) and it was my impression that they had a lot more to take in before production. Less proficient listeners apparently needed more time to process speech and identify NS. It is an issue for future investigation.

### 5.3 CONCLUDING WORDS

To conclude, it is important to raise awareness of the importance of making the results of previous studies, as well as of future research, available to teachers and textbook developers so that applied linguistics research really fulfills one of its objective, which is that of informing IL

pedagogy (Derwing & Munro, 2005). Finally, besides setting the primary phonological features to teach, teaching education needs to address, in undergraduate programs, the importance of explicit instruction of pronunciation and enable teachers-to-be with linguistic and methodological knowledge in order to implement the teaching of pronunciation in Brazil (Derwing & Munro, 2005). By joining forces, (1) well-prepared teachers, (2) pedagogy informed by research on intelligibility and IL-English acquisition, and (3) learners' commitment to IL-English learning should make the aim of successful communication attainable. Let us teach pronunciation, shall we? After all, what seems important is to build up learners' confidence about their own pronunciation.

## REFERENCES

Andrade, Melvin R. (2005). Performance of Japanese EFL Learners on a Test of English Syllabification, Word Stress, Sentence Focus, and Thought Groups. *Sophia Junior College Bulletin*, 25, Mar 2005, 1-25. Retrieved from

<https://www.jrc.sophia.ac.jp/pdf/research/bulletin/ki25/andrade.pdf>

Atechi, S. N. (2004). *The intelligibility of native and non-native English speech: A comparative analysis of Cameroon English and American and British English* (Doctoral Dissertation, Technischen Universität Chemnitz: Germany). Retrieved from

[http://www.qucosa.de/fileadmin/data/qucosa/documents/4843/data/atechi\\_intelligibility.pdf](http://www.qucosa.de/fileadmin/data/qucosa/documents/4843/data/atechi_intelligibility.pdf)

Bamgbose, A. (1998). Torn between the norms: innovations in world Englishes. *World Englishes*, 17(1), 1-14. doi: 10.1111/1467-971X.00078

Baptista, B. (2001). Frequent pronunciation errors of Brazilian learners of English. In M. Fortkamp & R. Xavier (Eds.), *EFL teaching and learning in Brazil: Theory and Practice*. Florianópolis: Insular.

Becker, M. R. (2011). A questão da inteligibilidade do inglês como língua franca. *Proceedings VII Congresso Internacional da Abralin Curitiba 2011* (pp. 2789–2800). Curitiba.

Becker, M.R. (2013). *Inteligibilidade da Língua Inglesa sob o paradigma de Língua Franca: Percepção de discursos de falantes de diferentes L1s por brasileiros* (Doctoral Dissertation, Universidade Federal do Paraná, Curitiba). Retrieved from <http://acervodigital.ufpr.br/bitstream/handle/1884/32270/R%20-%20T%20-%20MARCIA%20REGINA%20BECKER.pdf?sequence=1>

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., Schoenemann, T. (2009). Language Is a Complex Adaptive System: Position Paper. *Language Learning*, 59(1), 1-26.

In-text: (Beckner et al., 2009).

Berns, M. (2008). World Englishes, English as a lingua franca, and intelligibility. *World Englishes*, 27(3), 327-334.

Boersma, P., & Weenink, D. (2014). *Praat*: doing phonetics by computer [Computer program]. Version 5.4.04. Retrieved from <http://www.praat.org/>

Bolinger, D. (1998). Intonation in American English. In D. Hirst & A. Di Cristo (Eds.), *Intonation System: A survey of twenty languages* (pp. 45-55). Cambridge: Cambridge University Press.

Bolinger, D. (1972). Accent is predictable (if you are a mind reader). *Language*, 48(3). September 1972, 633-644. Retrieved from [https://www.jstor.org/stable/412039?seq=1#page\\_scan\\_tab\\_contents](https://www.jstor.org/stable/412039?seq=1#page_scan_tab_contents)

Bradlow, A. R. & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener- and item-related factors. *Journal of the Acoustical Society of America*, 106(4), 2074-2085.

Brawerman, A. (2006). *Uma análise de erros de estudantes brasileiros de inglês na acentuação de palavras com sufixos*. (Master Thesis, Universidade Federal do Paraná, Curitiba).

Büring, D. (2012). Focus and Intonation. In G. Russell & D. G. Fara (Eds.), *The Routledge Companion to the Philosophy of Language* (pp. 103-115). London: Routledge. Draft retrieved from <http://semanticsarchive.net/Archive/zFmZmM4N/Buring.2012.Focus%20and%20Intonation.RoutlegdePhilLgCompanion.pdf>

Calhoun, S. (2007). Predicting Focus through Prominence Structure. *Proceedings of Interspeech* (pp. 622-625). Antwerp, Belgium. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/publications/sasha-calhoun/calhounIS07.pdf>

Calhoun, S., Nissin, M., Steedman, M., & Brenier, J. (2005). A Framework for Annotating Information Structure in Discourse. *Proceedings of Frontiers in Corpus Annotation II: Pie in the Sky*,



ACL2005 Conference Workshop (pp. 45-42). Ann Arbor, Michigan.  
Retrieved from <http://groups.inf.ed.ac.uk/switchboard/infostruc.pdf>

Carpes, D. F. R. P. (2014). *Um estudo prosódico-semântico da não exaustividade no português brasileiro* (Master Thesis, Universidade Federal de Santa Catarina, Florianópolis, SC).

de Castro Gomes, M. L. (2013). Understanding the Brazilian way of speaking English. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, Aug. 2012. (pp. 279-289). Ames, IA: Iowa State University. Retrieved from <http://jlevis.public.iastate.edu/pslltconference/4th%20Proceedings/Castro%20Gomes%20PSLLT%202012.pdf>

Catford, J. (1950). Intelligibility. *English Language Teaching Journal*, 1(1), 7-15.

Celce-Murcia, M., Brinton, D.M., & Goodwin, J.M. (1996). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. New York, NY, CUP.

Chomsky, N., Halle, M. (1968). *The sound pattern of English*. New York, Harper & Row Publishers.

Crystal, D. (2008). *A Dictionary of Linguistics and Phonetics*. 6<sup>th</sup> Edition. Malden, MA: Blackwell Publishing.

Crystal, D. (2003). *English as a global language*. Cambridge: CUP.

Cruz, N. C. (2008). Familiaridade do ouvinte e inteligibilidade da pronúncia de aprendizes brasileiros de inglês. *Revista Horizontes de Lingüística Aplicada*, 7(1), 88-103. Retrieved from <http://periodicos.unb.br/index.php/horizontesla/article/view/2998/2600>

Cruz, N. C. (2004). *Pronunciation Intelligibility in Spontaneous Speech of Brazilian Learners' English* (Doctoral Dissertation, Universidade Federal de Santa Catarina, Florianópolis).

Cruz, N. C. (2003). An exploratory study of pronunciation intelligibility in the Brazilian Learner's English. *the ESspecialist*, 24(2), 155–175.

Retrieved from

[http://www.corpuslg.org/journals/the\\_especialist/issues/24\\_2\\_2003/AR\\_TIGO3.PDF](http://www.corpuslg.org/journals/the_especialist/issues/24_2_2003/AR_TIGO3.PDF)

Dalton, C., & Seidlhofer, B. (1995). *Pronunciation*. Oxford: Oxford University Press.

Derwing, T. M., & Munro, M. J. (1997). Accent, comprehensibility and intelligibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19, 1-16.

Derwing, T. M., & Munro, M. J. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, 39(3), 379-397.

Elordieta, G. (2007). Constraints on intonational prominence of focalized constituents. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation* (pp. 1-22). Heidelberg, New York, London: Springer. Retrieved from <http://gcp.ruhosting.nl/carlos/matt.pdf>

Fernandes-Svartman, F. R. (2012). A entoação das sentenças clivadas em português brasileiro e a interface sintaxe-fonologia. *Filologia e Linguística Portuguesa*, 14(1), 37-56. doi: <http://dx.doi.org/10.11606/issn.2176-9419.v14i1p37-56>

Fernandes-Svartman, F. R. (2008). A distinção foco/tópico em tzotzil, jakalteq, tembé, xavante, português brasileiro e português europeu. *Revista de Estudos da Linguagem*, 16(2), 55-87. doi: <http://dx.doi.org/10.17851/relin.v16i2.2493>

Fernandes, F. R. (2007). Tonal association in neutral and subject-narrow-focus sentences of Brazilian Portuguese: a comparison with European Portuguese. *Journal of Portuguese Linguistics*, 5/6, 91-115.

Field, J. (2003). The fuzzy notion of 'intelligibility': A headache for pronunciation teachers and oral testers. *IATEFL Special Interest Groups Newsletter*, 35-38.

Firth, A. (1996). The discursive accomplishment of normality: On conversation analysis and 'lingua franca' English. *Journal of Pragmatics*, 26(2), 237–259.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In: W. Strange (Ed.). *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Frota, S., MORAES, J. A. de (2016). Intonation in European and Brazilian Portuguese. In W. L. Wetzels, J. Costa & S. Menuzzi (Eds.), *The Handbook of Portuguese Linguistics* (pp. 141-166). Hoboken: Wiley-Blackwell.

Frota, S.; Cruz, M.; Svartman, F. R.; Collischonn, G.; Fonseca, A.; Serra, C.; Oliveira, P.; Vigário, M. Intonational variation in Portuguese: European and Brazilian varieties. In: S. Frota & P. Prieto (Eds.), *Intonational Variation in Romance* (pp. 235-283). Oxford: Oxford University Press. doi: <https://doi.org/10.1093/acprof:oso/9780199685332.003.0007>.  
In-Text: (Frota et al., 2015).

Galaczi, E., Post, B., Li, A., Barker, F., Schmidt, E. (2017). Assessing Second Language Pronunciation: Distinguishing Features of Rhythm in Learner Speech at Different Proficiency Levels. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 157-182). Bristol, UK: Multilingual Matters Ltd.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context and consequence. In H. Giles, N. Coupland & J. Coupland (Eds.), *Contexts of Accommodation: Developments in applied sociolinguistics* (pp. 1-68). Cambridge: CUP.

Gomes, M. L. de C., Brawerman-Albini, A., Engelbert, A. P. P. F. (2014). The perception of Vowel Epenthesis and Word Stress in an English as a Lingua Franca Context. *Proceedings of the International Symposium on the Acquisition of Second Language Speech Concordia*

*Working Papers in Applied Linguistics*, 5, ©2014 COPAL. Retrieved from [http://doe.concordia.ca/copal/documents/15\\_Gomes\\_etal\\_Vol5.pdf](http://doe.concordia.ca/copal/documents/15_Gomes_etal_Vol5.pdf)

Gonçalves, A. R. (2014). *In search of speech intelligibility: The case of English high front vowels* (Master Thesis, Universidade Federal de Santa Catarina, Florianópolis).

Gordon, M. (2014). Disentangling stress and pitch accent: Toward a typology of prominence at different prosodic levels. In H. van der Hulst (Ed.), *Word Stress: Theoretical and Typological Issues* (pp. 83-118). Oxford: Oxford University Press. Retrieved from [http://www.linguistics.ucsb.edu/faculty/gordon/Gordon\\_Stress&accent.pdf](http://www.linguistics.ucsb.edu/faculty/gordon/Gordon_Stress&accent.pdf)

Gussenhoven, Carlos (2015). Suprasegmentals. In James D. Wright (ed.) *International Encyclopedia of Social and Behavioral Sciences* Second Edition (pp. 714-721). Oxford: Elsevier. Retrieved from <http://gеп.ruhosting.nl/carlos/52024Suprasegmentals.pdf>

Gussenhoven, C. (2007). Types of Focus in English. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation* (pp. 83-100), Heidelberg, New York, London: Springer. Retrieved from <http://gеп.ruhosting.nl/carlos/matt.pdf>

Gussenhoven, Carlos (1994). Focus and sentence accents in English. In P. Bosch & R. van der Sandt (Eds.), *Focus & Natural Language Processing. Working Papers of the Institute for Logic and Linguistics* (pp. 83-92). IBM Deutschland Informationssysteme GmbH, Heidelberg. Retrieved from [http://gеп.ruhosting.nl/carlos/focus\\_and\\_sentence\\_accents.pdf](http://gеп.ruhosting.nl/carlos/focus_and_sentence_accents.pdf)

Gussenhoven, C. (1983). Stress shift and the nucleus. *Linguistics* 21, 303-339. Retrieved from [http://gеп.ruhosting.nl/carlos/Stress\\_shift\\_and\\_the\\_nucleus.pdf](http://gеп.ruhosting.nl/carlos/Stress_shift_and_the_nucleus.pdf)

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201-223.

- Harris, K. (2014). The Prosodic Marking of Givenness in English and Italian: A Comparative Study. *Newcastle and Northumbria Working Papers in Linguistics*, 20, 28-43. Retrieved from <http://www.ncl.ac.uk/media/wwwnclacuk/crills/files/vol21-harris.pdf>
- Hedberg, N., & Sosa, J. M. (2007). The prosody of topic and focus in spontaneous English dialogue. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation* (pp. 101-120). Heidelberg, New York, London: Springer.
- Heusinger, K. (2007). Discourse Structure and Intonational Phrasing. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and focus: Cross-linguistic perspectives on intonation and meaning* (pp. 265-290). Studies in Linguistics and Philosophy Series, Kluwer Academic Publishers. Retrieved <http://gep.ruhosting.nl/carlos/matt.pdf>
- Hirst, D. (1998). Intonation in British English. In D. Hirst & A. Di Cristo (Eds.) *Intonation System: A survey of twenty languages* (pp. 56-77). United Kingdom: Cambridge University Press.
- James, C. (1998). *Errors in language learning and use: exploring error analysis*. London: Longman.
- Jenkins, J. (1997). Teaching intonation for English as an International Language: teachability, learnability and intelligibility. *Speak Out! Newsletter of the IATEFL Pronunciation Special Interest Group*, 21, 15-26.
- Jenkins, J. (2000). *The phonology of English as an International Language*. Oxford: Oxford University Press.
- Kachru, B. (1986). The Power and Politics of English. *World Englishes*, 5(2-3), 113-282.
- Kachru, B. (1992). *The other tongue*. Urbana: University of Illinois Press.

Kachru, B., Kachru, Y. & Nelson, C. L. (2006). (Eds.). *The Handbook of World Englishes*. Oxford, London: Blackwell Publishing Ltd.

Keating, P., Baroni, M., Mattys, S., Scarborough, R., Alwan, A., Auer, E. T., Bernstein, L. E. (2003). Optical Phonetics and Visual Perception of Lexical and Phrasal Stress in English. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *15<sup>th</sup> International Congress of Phonetic Sciences Proceedings*, August 3-9 (pp. 2071-2074). Spain. Retrieved from [https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15\\_2071.pdf](https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_2071.pdf)  
In-Text: (Keating et al., 2003).

Kenworthy, J. (1987). *Teaching English pronunciation*. London: Longman.

Kirkpatrick, A. (2009). English as the international language of scholarship: Implications for the dissemination of ‘local’ knowledge. In F. Sharifian (Ed.) *English as an international language: Perspectives and pedagogical issues* (pp. 254-270). Bristol: Multilingual Matters Ltd.

Kirkpatrick, A. (2007). *World Englishes. Implications for International Communication and English Language Teaching*. Cambridge: Cambridge University Press.

Kivistö-de-Souza, H. (2017). Examining L1 Brazilian Portuguese speakers’ sensitivity to English nuclear stress assignment. *Revista de Estudos da Linguagem*, 25(2), 483-514. doi: <http://dx.doi.org/10.17851/relin.v25i2.10462>

Klok, V. J., Wagner, M., Goad, H. (2011). Contextual influence on acoustic correlates of focus in English, Québec French and European French. *Experimental and theoretical advances in prosody* (ETAP), McGill University, Montreal (poster). Retrieved from [http://www.prosodylab.org/~chael/www/etap2/abstracts/vanderklok\\_etal.pdf](http://www.prosodylab.org/~chael/www/etap2/abstracts/vanderklok_etal.pdf)

Krahmer, E., & Swerts, M. (2007). Perceiving Focus. In C. Lee, M. Gordon, & D. Büring (Eds.) *Topic and focus: Cross-linguistic*

*perspectives on intonation and meaning* (pp. 121-138). Studies in Linguistics and Philosophy Series, Kluwer Academic Publishers.

Krifka, M. (2007). The semantics of questions and the focusation of answers. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and focus: Cross-linguistic perspectives on intonation and meaning* (pp. 139-150). Studies in Linguistics and Philosophy Series, Kluwer Academic Publishers. Retrieved from <http://amor.cms.hu-berlin.de/~h2816i3x/QuestionsSantaBarbara.pdf>

Ladd, 1980. *The Structure of Intonational meaning: Evidence from English*. Indiana University Press, Bloomington, 50-99. Retrieved from <http://www.ling.upenn.edu/~kroch/courses/lx650/650-07/ladd80.pdf>

Lanham, L. W. (1984). Stress and Intonation and the intelligibility of Sound African Black English, *African Studies*, 43(2), 217-230. doi: 10.1080/00020188408707626.

Lambrecht, K. (1998). *Information structure and sentence form: topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18(2), 141-165.

Larsen-Freeman, D. (2007). Reflecting on the cognitive-social debate in second language acquisition. *The Modern Language Journal*, 91, 773-787.

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

Leffa, V. J. (2002). Teaching English as a multinational language. *The Linguistic Association of Korea Journal*, 10(1), 29-53.

Lege, R. F. (2012). *The Effect of Pause Duration on Intelligibility of Non-Native Spontaneous Oral Discourse* (Master Thesis, Brigham Young University, Provo, UT). Retrieved from <http://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=4487&context=etd>

Levis, J. (2011). Assessing speech intelligibility: Experts listen to two students. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2010. (56-69), Ames, IA: Iowa State University.

Liberman, M. Y. (1975). *The Intonational System of English* (Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge). Retrieved from <http://www.ai.mit.edu/projects/dm/theses/liberman75.pdf>

Lindemann, S. (2010). Who's "Unintelligible"? The Perceiver's Role. *Issues in Applied Linguistics*, 18(2), 223-232. Retrieved from <http://escholarship.org/uc/item/89f0w1ch>

Liu, F. (2010). Single vs double focus in English statements and yes/no questions. *Proceedings Speech Prosody 2010*. Chicago, Illinois. Retrieved from <http://speechprosody2010.illinois.edu/papers/100975.pdf>

Lucente, L. (2012). *Aspectos Dinâmicos da Fala e da Entoação no Português Brasileiro* (Doctoral Dissertation, Universidade Estadual de Campinas, Campinas).

Manosso, P. Z. (2013). *O emprego do acento nuclear em língua inglesa por falantes brasileiros: um estudo comparativo na era do inglês como língua franca* (Trabalho de conclusão de curso, Universidade Tecnológica Federal do Paraná, Curitiba). DACEX/DALEM, Curitiba, PR. Retrieved from [http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/1072/1/CT\\_COLET\\_2012\\_2\\_10.pdf](http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/1072/1/CT_COLET_2012_2_10.pdf)

Marta, E. de S. (2011). *Avaliação da produção e percepção da prosódia na expressão de modalidades por falantes de inglês como L1 e aprendizes de inglês como L2* (Doctoral Dissertation, Universidade Federal do Rio de Janeiro, Rio de Janeiro).

Mary, L. (2012). *Extraction and Representation of Prosody for Speaker, Speech and Language Recognition*. Springer Briefs in Speech Technology. Springer: New York. doi: 10.1007/978-1-4614-1159-8 1



Mateus, M. H. M. (2004). Estudando a melodia da fala: traços prosódicos e constituintes prosódicos. Palestra de abertura do Encontro sobre O Ensino das Línguas e a Linguística – APL e ESE de Setúbal, 27 e 28 de setembro. Retrieved from <http://www.iltec.pt/pdf/wpapers/2004-mhmateus-prosodia.pdf>

Matic, D. (2015). Information Structure in Linguistics. In J. D. Wright (Ed.), *International Encyclopedia of Social and Behavioral Sciences* (2nd edition). Oxford: Elsevier. Retrieved from [http://pubman.mpdl.mpg.de/pubman/item/escidoc:1703515:1/component/escidoc:2152081/Matic\\_information\\_2015.pdf](http://pubman.mpdl.mpg.de/pubman/item/escidoc:1703515:1/component/escidoc:2152081/Matic_information_2015.pdf)

McKay, S. (2002). *Teaching English as an international language: Rethinking goals and approaches*. Oxford: Oxford University Press.

McKay, S. L. (2010). English as an International Language. In N. H. Hornberger and S. L. McKay (Eds), *Sociolinguistics and Language Education: New Perspectives on Language & Education*. (pp. 89-115). Cleverdon: Multilingual Matters Ltd.

McKay, S. L. (2003). Toward an Appropriate EIL pedagogy: re-examining common ELT assumptions. *International Journal of Applied Linguistics*, 13(1), 1-22.

Mello, H., & Silva, L. F. L. (2015). Focus and information patterning: refining terminology and distinguishing categories in a spoken corpus. *ReVEL, edição especial n. 10*, 138-169. Retrieved from <http://revel.inf.br/files/509515df4886ec4107a14de1eb603e8b.pdf>

Migliorini, L.; & Massini-Cagliari, G. (2010). Sobre o ritmo do Português Brasileiro: evidências de um padrão acentual. *ReVEL*, 8(15). Retrieved from [http://www.revel.inf.br/files/artigos/revel\\_15\\_sobre\\_o\\_ritmo.pdf](http://www.revel.inf.br/files/artigos/revel_15_sobre_o_ritmo.pdf)

Mioto, C. (2003). Focalização e Quantificação. *Revista Letras*, 61(especial), 169-189. Curitiba: Editora UFPR.

- Moraes, J. A. (1998). Intonation in Brazilian Portuguese. In: Hirst, D.; DI Cristo, A. (Ed.). *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press, p. 179-194. doi: <https://doi.org/10.1080/09658416.2011.609621>.
- Moraes, J. A. de (2008). The Pitch Accents in Brazilian Portuguese: analysis by synthesis. *Speech Prosody*, 20, May 6-9, pp. 1-9.
- Moraes, J. A. (2007). Intonational phonology of Brazilian Portuguese. Poster session presented at the *Workshop on Intonational Phonology, 16<sup>th</sup> International Congress of Phonetic Sciences*, 2007, Saarbrücken.
- Moraes, J. A., Carnaval, M., Coelho, A. B. B. (2015). A manifestação prosódica do foco em interrogativas totais no Português do Brasil e sua percepção. *ReVEL, 10*(especial), 170-194. Retrieved from <http://revel.inf.br/files/25628f323ed484f9952532a1604fbb93.pdf>
- Moyer, A. (2013). *Foreign Accent: The phenomenon of Non-Native Speech*. New York: Cambridge University Press.
- MOTU Ultra Lite mk3 [a hybrid FireWire USB2 audio interface for Mac and Windows]. *MOTU*, Cambridge, MA.
- Munro, M. J., & Bohn, O.-S. (2007). The study of second language speech: A brief overview. In O.-S. Bohn, *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. (pp. 3-12). Amsterdam, NLD, John Benjamins Publishing Company.
- Na, H., Yuan, J., Bin, L. (2012). Phonetic and Phonological Realization of Narrow Focus in English Declarative Sentences by Zhenjiang EFL Learners. *Proceedings Speech Prosody, 6<sup>th</sup> International Conference*, Shanghai, China, May 22-25. Retrieved from [http://isle.illinois.edu/sprosig/sp2012/uploadfiles/file/sp2012\\_submission\\_205.pdf](http://isle.illinois.edu/sprosig/sp2012/uploadfiles/file/sp2012_submission_205.pdf)
- Nava, E., & Zubizarreta, M. L. (2010). Deconstructing the Nuclear Stress Algorithm: Evidence from Second Language Speech. In N. Erteschik-Shir & L. Rochman (Eds.), *The Sound Patterns of Syntax* (pp. 291-316). Oxford: Oxford University Press.

Nava, E., & Zubizarreta, M. L. (2008). Prosodic Transfer in L2 Speech: Evidence from Phrasal Prominence and Rhythm. *Speech Prosody*, May 6-9, Campinas, Brazil, 335-338. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.537.7561&rep=rep1&type=pdf>

Nelson, C. L. (2008). Intelligibility since 1969. *World Englishes*, 27(3/4), 297-308. doi: 10.1111/j.1467-971X.2008.00568.x

Nooteboom, S. (1997). The prosody of speech: melody and rhythm. In (Ed.), *The Handbook of Phonetic Sciences* (pp. 640-673). Blackwell Handbooks in Linguistics. Draft retrieved from <https://www.phil.uu.nl/tst/2012/Werk/PROSTEM.pdf>

OcenAudio [multiplatform audio editor]. (2015). Retrieved from <http://www.ocenaudio.com>

Olson, H. F. (1972). The Mesurement of Loudness. *Audio, February*, 18-22.

O'Mahony, J. (2014). *The Prosodic Cues Used in Disambiguation of Disjunctive Questions* (Master Thesis, University of Amsterdam, Amsterdam. Retrieved from <http://www.fon.hum.uva.nl/theses/JohannahOMahonyBA2014.pdf>

Ortiz-Lira, H. (1998). Word stress and sentence accent. *Cuadernos de la Facultad, Monografías Temáticas, 16*, Faculdade de Historia, Geografia y Letras, Universidad Metropolitana de Ciencias de La Educación, Santiago de Chile, mayo de 1998. Retrieved from <https://mariaemramamos.wikispaces.com/file/view/Word+Stress+and+Sentence+Accent.pdf>

Passarella-Reis, L. (2014). Descendo ou rising? Padrões entoacionais de perguntas totais da língua inglesa e da língua portuguesa por falantes brasileiros. *Revista X, 1*, 1-22. doi: <http://dx.doi.org/10.5380/rvx.v1i0.34266>

Passarella-Reis, L., Gonçalves, A. R., Silveira, R. (2016). Perception of Intonational Patterns and Speaker's Intentionality in English Yes-No Questions Produced by Brazilians. *RELIN*, 24(1), 65-98. doi: 10.17851/2237-2083.24.1.65-97

Passarella-Reis, L., & Silveira, R. (2016). Nuclear stress placement by Brazilian Users of English as an International Language. *Domínios de Lingu@agem*, 10(2), 673-702. doi: <http://dx.doi.org/10.14393/DL22v10n2a201611>

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English Intonation* (Doctoral Dissertation, Department of Linguistics and Philosophy, MIT, Cambridge, MA). Retrieved from [http://faculty.wcas.northwestern.edu/~jbp/publications/Pierrehumbert\\_PD.pdf](http://faculty.wcas.northwestern.edu/~jbp/publications/Pierrehumbert_PD.pdf)

Pierrehumbert, J., Hirschberg, J. (1992). The Meaning of Intonational Contours in the Interpretation of Discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in Communication* (2<sup>nd</sup> edition) (pp. 271-311). Cambridge, Massachusetts: MIT Press.

Quarezemin, S., Neto, A. T. (2015). Da sintatização dos focos contrastivo e exaustivo em CP e das estratégias de marcação de foco. *ReVEL*, 10(especial), 42-77. Retrieved from <http://www.revel.inf.br/files/eb575603b243ddb4990470247c2ece43.pdf>

Ribeiro, N. L. (2015) Processamento do foco prosódico em clivadas invertidas reduzidas. *ReVEL*, 10(especial), 317-341. Retrieved from <http://revel.inf.br/files/38b012b7562bb6e51b8d33f28e66e03f.pdf>

Roach, P. (2009). *English Phonetics and Phonology* (4<sup>th</sup> edition). Cambridge: CUP.

Roelofsen, F., Gool, S. van (2009). Disjunctive questions, intonation, and highlighting. *Proceedings of the 17<sup>th</sup> Amsterdam Colloquium Conference on Logic, Language, and Meaning*, Springer-Verlag, Berlin, Heidelberg © 2010, 384-394. Retrieved from <https://www.ilc.uva.nl/Research/Publications/Reports/PP-2010-09.text.pdf>

Schadech, T.S. (2013). *The production of word-initial /ɹ/ by Brazilian learners of English and the issues of comprehensibility and intelligibility* (Master Thesis, Universidade Federal de Santa Catarina. Florianópolis, SC). Retrieved from <https://repositorio.ufsc.br/bitstream/handle/123456789/103510/316556.pdf?sequence=1>

Schadech, T. S; & Silveira, R. (2013). How do the non-target pronunciations of the consonants /θ/ and /ð/ by Brazilian learners of English affect comprehensibility? *Cadernos do IL*, 46, 4-23. Retrieved from <http://seer.ufrgs.br/index.php/cadernosdoil/article/view/35862/25715>

Seidlhofer, B. (2009). Accommodation and the idiom principle in English as a Lingua Franca. *Intercultural Pragmatics*, 6(2), 195-215. doi: 10.1515/IPRG.2009.011

Seidlhofer, B. (2008). Introducing English as a lingua franca (ELF): Precursor and partner in intercultural communication. *Synergies Europe*, 3, 25-36. Retrieved from <https://gerflint.fr/Base/Europe3/hulmbauer.pdf>

Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, 59, 339-341. Retrieved from [https://docs.ufpr.br/~clarissa/pdfs/ELF\\_Seidlhofer2005.pdf](https://docs.ufpr.br/~clarissa/pdfs/ELF_Seidlhofer2005.pdf)

Sharifian, F. (2009). *English as an International Language: Perspectives and Pedagogical Issues*. Bristol: Multilingual Matters.

Shue, Y., Iseli, M., Veilleux, N., & Alwan, A. (2007). Pitch Accent versus Lexical Stress: Quantifying Acoustic Measures Related to the Voice Source. *Proceedings of Interspeech 2007* (pp. 2625-2628). Antwerp, Belgium. Retrieved from [http://www.seas.ucla.edu/spapl/paper/iseli\\_interspeech\\_2007.pdf](http://www.seas.ucla.edu/spapl/paper/iseli_interspeech_2007.pdf)

Silveira, R. (2011). Avaliando a proficiência oral em língua inglesa: o papel dos avaliadores e dos informantes. In: M. B. Mota, J. M. Bergsleithner, & J. Weissheimer (Eds.), *Produção Oral em LE: Múltiplas Perspectivas*, 19 (pp. 73-96). Campinas: Pontes.

Sluijter, A. M. C., & van Heuven, V. J. (1996a). Spectral balance as an acoustic correlate of linguistic stress. *J. Acoustical Society of America*, 100(4), 2471-2485. Retrieved from [https://openaccess.leidenuniv.nl/bitstream/handle/1887/11173/5\\_167\\_118.pdf?sequence=1](https://openaccess.leidenuniv.nl/bitstream/handle/1887/11173/5_167_118.pdf?sequence=1)

Sluijter, A. M. C., & van Heuven, V. J. (1996b). Acoustic correlates of linguistic stress and accent in Dutch and American English. *International Conference on Spoken Language, ICSLP, Proceedings – ICSLP* (pp. 630-633). Retrieved from <http://www.asel.udel.edu/icslp/cdrom/vol2/604/a604.pdf>

Sluijter, A. M. C., & van Heuven, V. J. (1995). Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch, *Phonetica*, 52, 71–89.

Smith, L. E. (1988). Spread of English and issues of intelligibility. In Peter H. Lowenberg (Ed.) *Georgetown University Round Table on Language and Linguistics 1987 - Language Spread and Language Policy: Issues, Implications, and Case Studies* (pp. 265-282). Washington, D.C., Georgetown University Press. Retrieved from [https://repository.library.georgetown.edu/bitstream/handle/10822/555480/GURT\\_1987.pdf?sequence=1&isAllowed=y](https://repository.library.georgetown.edu/bitstream/handle/10822/555480/GURT_1987.pdf?sequence=1&isAllowed=y)

Smith, L. E., & Nelson, C.L. (1985). International intelligibility of English: directions and resources. *World Englishes*, 4, 333–342. doi: 10.1111/j.1467-971X.1985.tb00423.x

Smith, L. E., & Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly*, 13(3), 371-380. doi: 10.2307/3585884

Sosa, J. M., Seara, I. C., Nunes, V. G. (2013). Variação prosódica das sentenças interrogativas totais no falar catarinense: um estudo experimental. *Revista Leitura*, 2(52), 2 semestre/2013, 139-163. Retrieved from <http://www.seer.ufal.br/index.php/revistaleitura/article/view/1477/1005>

Tiffen, B. (1974). *The Intelligibility of Nigerian English* (Doctoral Dissertation, University of London, London). Retrieved from <http://discovery.ucl.ac.uk/1349554/1/475219.pdf>

t'Hart, J. (1981). Differential sensitivity to pitch distance, particularly in speech. *Journal of Acoustical Society of America*, 69(3), 811-821. Retrieved from <http://alexandria.tue.nl/repository/freearticles/734583.pdf>

Trofimovich, P., Kennedy, S. (2014). Interactive alignment between bilingual interlocutors: Evidence from two information-exchange tasks. *Bilingualism: Language and Cognition*, 17, 1-15.

Truckenbrodt, H., Sandalo, F., Abaurre, M. B. (2008). Elements of Brazilian Portuguese intonation. *Journal of Portuguese Linguistics*, 8-1 (5), 77-115.

Warren, P. (2013). *Introducing Psycholinguistics*. CUP, New York.

Widdowson, H. G. 1994. The ownership of English. *TESOL Quarterly*, 28, 377-389. doi: 10.2307/3587438

Wijaya, D. T., Yeniterzi, R. (2011). Understanding Semantic Change of Words Over Centuries. *DETECT 2011 - Proceedings of the 2011 international workshop on DETecting and Exploiting Cultural diversity on the social web* (pp. 35-40). Glaskow, Scotland, UK. Retrieved from <http://rtw.ml.cmu.edu/papers/wijaya-detect11.pdf>

Yavas, M. (2011). *Applied English Phonology* (2<sup>nd</sup> edition). Malden, Ma: Blackwell Pub.

Zielinski, B. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, 36, 69-84. doi: <https://doi.org/10.1016/j.system.2007.11.004>

Zoghbor, W. S. (2010). *The Effectiveness of the Lingua Franca Core (LFC) in Improving the Perceived Intelligibility and Perceived Comprehensibility of Arab Learners at Post-Secondary Level* (Doctoral

Dissertation, University of Leicester, Leicester). Retrieved from <https://lra.le.ac.uk/bitstream/2381/9635/1/2011ZOGHBORWSEDD.pdf>

Zubizarreta, M. L. (2016). "Information Structure and Nuclear Stress." In C. Féry & S. Ishihara (Eds.), *The Oxford Handbook of Information Structure* (pp. 165-184). Oxford: Oxford University Press. Draft retrieved from [https://dornsife.usc.edu/assets/sites/295/docs/OUP.Nuclear\\_Stress\\_IF.pdf](https://dornsife.usc.edu/assets/sites/295/docs/OUP.Nuclear_Stress_IF.pdf)

Zubizarreta, M. L.; & Nava, E. (2011). Encoding discourse-based meaning: Prosody vs. syntax: Implications for second language acquisition. *Lingua*, 121(4), 652-669, 2011. doi: <https://doi.org/10.1016/j.lingua.2010.06.013>

Zuengler, J. (1991). Accommodation in native-nonnative interactions: Going beyond the “what” to the “why” in second language research. In H. Giles, N. Coupland, & J. Coupland (eds.), *Contexts of Accommodation* (pp. 223-244). New York: CUP.



## **APPENDICES**



## Appendix A: Invitation do participate addressed to Brazilian Speakers of English

### UNIVERSIDADE FEDERAL DE SANTA CATARINA PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS

Dear Student,

I am a PhD candidate at UFSC. For my study, I need intermediate level students as participants and I would like to invite you to be one.

Your participation in this phase of data collection consists in describing some slides, which may take you up to fifteen minutes.

Your help you bring in weight and success to my study. If you would like to participate, please write your name, phone number and email address in the chart below and I will contact you as soon as possible.

Name	phone number	email address

Now, check the days and the time you would be available for data collection.

Time	Mondays	Tuesdays	Wednesdays	Thursdays	Fridays
8:00 A.M.					
9:00 A.M.					
10:00 A.M.					
11:00 A.M.					
12:00 P.M.					
1:00 P.M.					
2:00 P.M.					
3:00 P.M.					
4:00 P.M.					
5:00 P.M.					
6:00 P.M.					

Other: \_\_\_\_\_

Thank you very much.

Leonice Passarella dos Reis  
PhD Candidate

**Appendix B: Invitation do participate addressed to Speakers of English from a nationality other than Brazilian**

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS

Dear Student,

I am a PhD candidate at UFSC. For my study, I need intermediate level students as participants and I would like to invite you to be one.

Your participation in this phase of data collection consists in describing some slides, which may take you up to fifteen minutes.

Your help will bring in weight and success to my study. If you would like to participate, please write your name, phone number and email address in the chart below and I will contact you as soon as possible.

Name	phone number	email address

Now, check the days and the time you would be available for data collection.

Time	Mondays	Tuesdays	Wednesdays	Thursdays	Fridays
8:00 A.M.					
9:00 A.M.					
10:00 A.M.					
11:00 A.M.					
12:00 P.M.					
1:00 P.M.					
2:00 P.M.					
3:00 P.M.					
4:00 P.M.					
5:00 P.M.					
6:00 P.M.					

Other:

---

Thank you very much.

Leonice Passarella dos Reis  
PhD Candidate

## Appendix C: Consent Term - Speakers

### UNIVERSIDADE FEDERAL DE SANTA CATARINA PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS: ESTUDOS LINGUÍSTICOS E LITERÁRIOS

#### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO Falantes

Sou **Leonice Passarella dos Reis**, estudante de doutorado da **Universidade Federal de Santa Catarina**. Faço pesquisa na área da fonética e fonologia e sou orientada pela professora doutora **Rosane Silveira**.

Convidamos você para ser participante da pesquisa: “**Como assim? O acento nuclear em inglês como língua internacional: Usos e interpretações**”. Essa pesquisa está sendo realizada para investigar a comunicação oral de falantes brasileiros de inglês e, através de seus resultados, trazer contribuições para o ensino de inglês no Brasil, haja vista ter um potencial para informar a pedagogia da área e editores de livros didáticos a respeito do que incluir no currículo de ensino da língua. Além disso, esta pesquisa pretende contribuir para esclarecimentos de assuntos ainda obscuros na literatura de fonética e fonologia.

Se aceitar fazer parte deste estudo, **você participará de duas etapas de coleta de dados**. Na primeira etapa, você (1) assinará o presente termo de consentimento para assegurar seus direitos por conta do uso de seus dados coletados, (2) responderá a um questionário e (3) participará da gravação em áudio de uma atividade de produção oral. Isso levará em torno de 15 (quinze) minutos. Na segunda etapa, você participará de quatro atividades em duplas com um falante de inglês de outra nacionalidade, que serão filmadas e gravadas em áudio, e realizará duas atividades escritas relacionadas a vocabulário. Isso levará em torno de 30 (trinta) minutos. As datas de realização dessas etapas serão combinadas com antecedência de acordo com a sua disponibilidade para realizá-las.

As atividades são muito parecidas com as que você já realiza nas suas aulas de inglês, por isso **os riscos são mínimos**. Você poderá ficar um pouco ansioso, sentimento comum quando alguém faz algo diferente do que é sua rotina. Para evitar que você se sinta assim, você receberá instruções detalhadas, em todas as etapas você terá o meu acompanhamento e quaisquer dúvidas ou dificuldades serão atendidas prontamente. Você poderá também sentir um possível cansaço mental. Para minimizar esse efeito, você poderá escolher fazer pequenos intervalos entre uma atividade e outra. A sua participação lhe dará a

oportunidade (1) de ter o seu nível de proficiência oral em inglês medido gratuitamente por profissionais de inglês, (2) de praticar o uso do inglês tanto na descrição das imagens quanto nas interações com outros falantes, (3) de tomar ciência a respeito do uso do acento nuclear, pouco enfatizado em sala de aula, e (4) de contribuir para a reflexão de quais assuntos relacionados à pronúncia do inglês incluir no currículo de inglês no Brasil.

As informações que você oferecer e o material gravado serão absolutamente confidenciais. **A sua identidade nunca será revelada.** Os participantes serão identificados como “Participante 1, Participante 2” e assim por diante. Caso deseje, após a coleta e a análise dos dados, podemos informar a você quais foram os resultados obtidos através desta pesquisa.

A sua participação é totalmente voluntária. Esclarecemos que não haverá compensação financeira em função da sua participação na pesquisa. Se porventura você tiver, por mínimas que sejam, despesas decorrentes de sua participação nesta pesquisa, as pesquisadoras garantem o ressarcimento dos gastos, bem como indenização diante de eventuais danos oriundos também da pesquisa. É possível, a qualquer momento, desistir de sua participação, informando às pesquisadoras de sua decisão, de modo que seus dados sejam desconsiderados. A coleta acontecerá na própria Universidade Federal de Santa Catarina, no Laboratório de Linguística Aplicada, localizado na sala 505, Prédio B, do CCE.

Esta pesquisa está de acordo com os critérios estabelecidos na Resolução 466/2012 e seus documentos complementares. Sua participação não acarreta nem prejuízos nem privilégios. Se tiver dúvidas e precisar de esclarecimentos adicionais, você deve procurar as pesquisadoras Leonice Passarella dos Reis ([leonicepr@gmail.com](mailto:leonicepr@gmail.com) / 48-8404-5761 ou na rua Virgilino Ferreira de Souza, 206, bloco I, ap. 1005, Barreiros, São José – CEP 88117-700) e Rosane Silveira ([rosane@cce.ufsc.br](mailto:rosane@cce.ufsc.br)). Caso considere necessário, você pode entrar em contato com o Comitê de Ética em Pesquisa com Seres Humanos (CEPSH) da Universidade Federal de Santa Catarina, localizado no Prédio Reitoria II (Edifício Santa Clara), R: Desembargador Vitor Lima, nº 222, sala 401, Trindade, Florianópolis/SC. CEP 88.040-400, e-mail: [cep.propesq@contato.ufsc.br](mailto:cep.propesq@contato.ufsc.br). Telefone: +55 48 3721-6094 ou estabelecer contato eletrônico através do endereço: <http://cep.ufsc.br/contato/>.

### **Consentimento Livre e Esclarecido**

Eu, \_\_\_\_\_,  
CPF nº \_\_\_\_\_, após ter recebido todos os esclarecimentos e ciente dos meus direitos, concordo em participar desta pesquisa, bem como autorizo a divulgação e a publicação de toda informação

por mim transmitida, exceto dados pessoais, em publicações e eventos de caráter científico. Desta forma, assino este termo, juntamente com a pesquisadora, em duas vias de igual teor, ficando uma via sob meu poder e outra em poder das pesquisadoras.

Florianópolis, \_\_\_\_ / \_\_\_\_ / \_\_\_\_\_

---

Leonice Passarella dos Reis  
Pesquisadora

---

Assinatura do Participante

## Appendix D: Consent Term - Listeners

### UNIVERSIDADE FEDERAL DE SANTA CATARINA PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS: ESTUDOS LINGUÍSTICOS E LITERÁRIOS

#### CONSENT TERM Listeners

My name is **Leonice Passarella dos Reis** and I am a PhD candidate at the English Graduate Program at **Federal University of Santa Catarina**. Currently, I am conducting a study on the field of phonetics and phonology and my advisor is Professor **Rosane Silveira**.

We would like to invite you to participate in the study: “**What do you mean? Nuclear stress placement in English as an International Language: Uses and Interpretations**”. This study aims to investigate Brazilians’ oral communication in English and bring in contributions to the teaching of English in Brazil, given that it has the potential to inform English pedagogy and textbooks editors with regard to the teaching syllabus. Moreover, this study intend to shed light on unsolved issues in phonetics and phonology literature.

If you accept to be a participant, **you will participate in two phases of data collection**. In the first phase, you will (1) sign the present Consent Term in order to assure your rights related to this data collection, (2) answer a questionnaire, and (3) perform an audio-recorded speech production task. This participation may take you about 15 (fifteen) minutes. In the second phase, you will engage in four pair-work activities together with a Brazilian speaker of English. These interactions will be both video and audio-recorded. In this phase, you will also perform two written activities related to vocabulary and you will be interviewed by the researcher. It will take you about 50 (fifty) minutes to participate in this phase of data collection. The days and the time for data collection will be previously set according to your availability to participate.

The activities that you will perform are very similar to the ones that you already have in your language classes. Therefore, **the risks are few**. You can feel a little anxious, which is a common feeling when someone joins in activities that escape their daily routine. In order to avoid that you feel this way, you will receive detailed instructions, in all phases you will have my company, and any questions or difficulties will be promptly answered/solved. You may also feel a possible mental tiring. In order to minimize this effect, you can choose to take little breaks in between tasks. Your participation will allow you to (1) have your



proficiency level in English rated for free by professionals in the English area, (2) practice your English both when describing the slides and when interacting with other speakers, (3) become aware of the use of nuclear stress in English, which is usually not emphasized in classes, and (4) contribute to the reflection about the pronunciation features in English that should be included in the English teaching syllabus in Brazil.

The information you will offer and the recorded material will be absolutely confidential. **Your identity will never be revealed.** Participants will be identified as “Participant 1, Participant 2” and so forth. In case you wish to have access to the results of this study, they will be made available for you.

Your participation is not mandatory. We would like to clarify that there will be no financial benefits related to your participation. If you happen to have any expenses because of your participation in this study, the researchers involved guarantee that they will pay them back to you and the researchers will indemnify you for any loss or damage caused by your participation in this study. You can, at any moment, give up on your participation by letting the researchers know about your decision, so that your data may be discarded. The data collection will take place at the Federal University of Santa Catarina itself, at the Applied Linguistics Laboratory (FONAPLI), located at Building CCE-B, room 505.

The present study meets the criteria set at Resolution 466/2012 and its complementary documents. Your participation will neither bring in benefits nor disadvantages to you. In case you have incoming questions regarding the development of the present study, you can contact the researchers Leonice Passarella dos Reis ([leonicepr@gmail.com](mailto:leonicepr@gmail.com) / 48-8404-5761 or at “rua Virgilino Ferreira de Souza, 206, bloco 1, ap. 1005, Barreiros, São José” – ZIP CODE 88117-700) and Rosane Silveira ([rosane@cce.ufsc.br](mailto:rosane@cce.ufsc.br)). If you think it is necessary, you can get in touch with the Ethics Board on Research involving human beings (CEPSH) at UFSC, located at Prédio Reitoria II (Edifício Santa Clara), R: Desembargador Vítor Lima, nº 222, room 401, Trindade, Florianópolis/SC. ZIP CODE 88.040-400, e-mail: [cep.propesq@contato.ufsc.br](mailto:cep.propesq@contato.ufsc.br). Phone number: +55 48 3721-6094 or contact them at <http://cep.ufsc.br/contato/>.

### Consent Term

I, \_\_\_\_\_,  
 passport number \_\_\_\_\_, after having received the  
 information above and being aware of my rights, agree to participate in this  
 study and authorize the publication of the information that I will provide, except  
 personal data, in scientific contexts. Being this so, I sign this term, together with

the researcher, in two pages of equal value, one of which will be kept under my possession.

Florianópolis, \_\_\_\_ / \_\_\_\_ / \_\_\_\_\_

---

Leonice Passarella dos Reis  
Researcher

---

Participant's signature

## Appendix E: Questionnaire

*Thanks for volunteering to participate in the present study. Please, answer the questionnaire. In case you have doubts, talk to the researcher.*

*Leonice Passarella dos Reis  
Researcher*

1. Name \_\_\_\_\_
2. Email address \_\_\_\_\_
3. Age \_\_\_\_\_
4. Nationality \_\_\_\_\_
5. First language \_\_\_\_\_
6. Level of education \_\_\_\_\_
7. Occupation \_\_\_\_\_
8. Where have you lived most of your life? (city and country)  
\_\_\_\_\_
9. Do you go to school/College? If so, explain.  
a.  Yes. \_\_\_\_\_  
b.  No. \_\_\_\_\_
10. How old were you when you first started having English classes?  
\_\_\_\_\_
11. How long have you been having English classes? (years)  
\_\_\_\_\_
12. Are you currently attending English classes?  Yes  No  
If so, where? \_\_\_\_\_  
How many hours per week? \_\_\_\_\_
13. According to your opinion, check the proficiency level that matches your oral proficiency in English.  
a.  basic  
b.  pre-intermediate  
c.  intermediate  
d.  pre-advanced  
e.  advanced  
f.  English-L1 like
14. Regarding the use of English, in average how many hours per week do you:  
a. speak? \_\_\_\_\_  
b. listen (to music, videos, etc)? \_\_\_\_\_  
c. write? \_\_\_\_\_  
d. read? \_\_\_\_\_

*Thanks! ☺*

**Appendix F: Oral Production Task**

**INSTRUCTIONS**

You will see 4 slides containing pictures.



Your task is to DESCRIBE each picture, by providing as much detail as possible.

You DON'T need to make connections between the pictures in the same slide.

If it helps you, try to answer the following questions:

- 1) What do you see in the picture?
- 2) What is happening in the picture?

## Example

Pictures	Possible description
 An illustration of four people in an office setting. Three people are seated around a table, and one person is standing, holding a piece of paper and a pen. There is a potted plant in the background.	<p>There's an office. There are four people discussing something. They are wearing formal clothes. One of them is holding a piece of paper and a pen. There is a plant...</p>
 A photograph of a garden path. The path is made of dirt and is surrounded by lush green foliage and many colorful flowers, including pink and purple ones. The scene is bright and sunny.	<p>The other picture shows a garden. There are lots of flowers. It's a beautiful sunny day...</p>

## Testing the recording equipment

- Before you begin the task, put on your headphone and record the following sentence:

**This task is interesting.**

(**Stop** and wait until the researcher checks if your voice was properly recorded.)

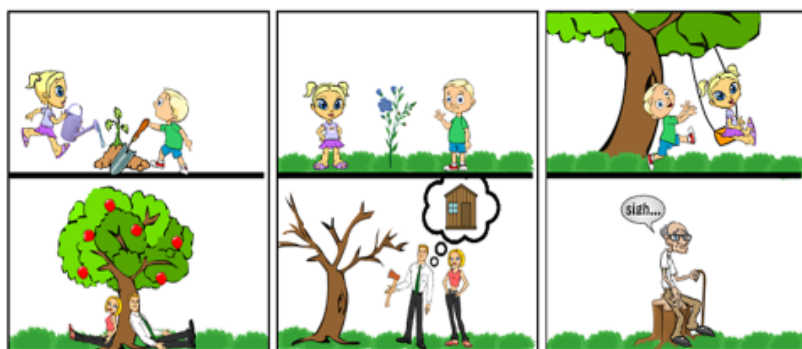
When you finish the description of a slide push ENTER to advance to the next slide.

**BEGIN THE TASK.**









The end.

Thank you!

## Appendix G: Instructions for the raters

### INSTRUCTIONS AND FORMS FOR RATERS<sup>78</sup>

The data you will work with come from a study in which English speakers with different proficiency levels were asked to describe several slides in English. Your task is to **evaluate their oral proficiency level** based on their oral descriptions of the slides on pages 4-6. Although they described 4 slides, the production of only one slide was used. Note that the participants may be describing the whole slide or only part of it, depending on how much they spoke. Participants who are very brief in their descriptions are describing the whole slide, while participants who speak a lot are describing only some of it. In order to guide their descriptions, they could answer two questions: What do you see in the picture? What are people doing?

There is no “correct” answer for this task. The amount of details provided by each participant depended not only on their proficiency level but also on their personalities (some people like providing more details than others). You are advised to wear earphones while listening to the descriptions.

#### **Steps for rating the participants:**

**STEP 1:** Listen to each participant speak for about 25-30 seconds and rate the participants’ oral proficiency level. To listen to the participants, open the audio file (Proficiency\_test.mp3).

**STEP 2:** Rate the participant’s oral proficiency level. To do this, you will have a scale like the one below for each participant. You should check (✓) the number that best defines the participant’s oral proficiency level, ranging from 1 (very low proficiency) to 10 (L1-like proficiency).

---

<sup>78</sup> Adapted from Silveira, 2011.

**Participant A****Form A: Oral proficiency**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency											L1-like proficiency

To make sure you understand the procedures, rate the performance of **Participant A**.

**Step 1:** Open the audio file for Participant A (2015\_Participant\_A.mp3).

**Step 2:** Using Form A, provided on the top of this page, rate the participant's performance in terms of oral proficiency level in English.

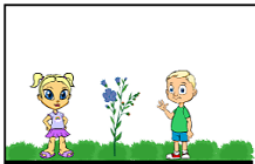
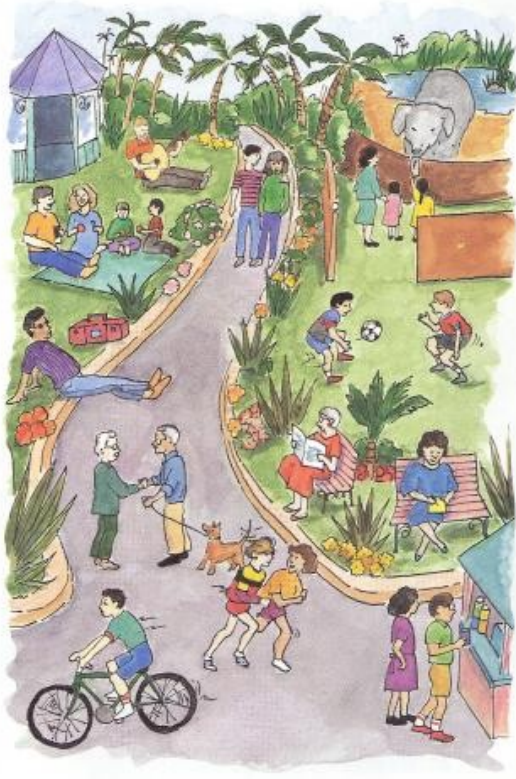
- ⇒ If you experienced any difficulty when rating Participant\_A, feel free to contact the researcher (leonicepr@gmail.com) to ask any questions about the procedures.
- ⇒ If you are ready, you can start rating the 33 participants. Clean copies with the forms and the names of each participant are in the Word file called Forms\_raters.doc. Remember saving the form every time you type in new information. When you are finished rating all participants, send me the file with the forms updated with your ratings for all participants.

Thank you!

Leonice Passarella dos Reis

Images described by the participants. Remember: participants are describing only one slide or part of it.









## Appendix H: Forms for the raters

Open the audio file (Proficiency\_test.mp3). The samples are ordered as follows: Participant 1, Participant 2, Participant 3 and so on. After listening to one participant's description, fill out that participant's form. In case you think it is necessary, pause the audio after each participant's description so you can fill out the forms.

### Participant 1

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

### Participant 2

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

### Participant 3

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

### Participant 4

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 5**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 6**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 7**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 8**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 9**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 10**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 11**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 12**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 13**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 14**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 15**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 16**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 17**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 18**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 19**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 20**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level

**Participant 21**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level

**Participant 22**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level

**Participant 23**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level

**Participant 24**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level

**Participant 25**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 26**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 27**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 28**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 29**

	0	1	2	3	4	5	6	7	8	9	10
	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level											L1-like proficiency level

**Participant 30**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level

**Participant 31**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level

**Participant 32**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level

**Participant 33**

0	1	2	3	4	5	6	7	8	9	10
( )	( )	( )	( )	( )	( )	( )	( )	( )	( )	( )
Very low proficiency level										L1-like proficiency level



## Appendix I: Task 1 – Breaking the ice



You work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English.

For this task, you will meet a coworker for the first time. You don't know the person's name, nationality, age or any further general information.

Try to get to know this coworker. Ask questions that you would normally ask when meeting someone. Try to sound as natural as possible.

Additionally, try to find out three things that you have in common. For example, a dislike of jilo, Brazilian country music, and sushi.

Remember: You need to use English.

Three things we have in common:

- 1) \_\_\_\_\_
- 2) \_\_\_\_\_
- 3) \_\_\_\_\_

## Appendix J: Task 2 – I know the answer! (Speaker’s version)



Remember that you work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English. For this task, you will consider different context questions, by reading them silently. Each question establishes a context that requests you to say a piece of information that your coworker needs. You will say aloud

that piece of information. To illustrate, see an example below.

(Talking to a coworker, who is a great fan of Adele.)

Read it silently: When will Adele be in Canada?

Say it aloud: **ADELE WILL BE IN CANADA IN DECEMBER.**

In the example above, you need to have in mind that the question you are answering is “When will Adele be in Canada?” as if your coworker had asked it. When you say the answer, you emphasize **DECEMBER**, because that’s the information your coworker was looking for.

**START**

**01)** (Talking about the way Zenny feels about your coworker.)

Read it silently: How does Zenny feel about me?

Say it aloud: **ZENNY LOVES YOU.**

**02)** (Your coworker has received some pearls from a secret admirer.)

Read it silently: Who sent the pearls to me?

Say it aloud: **ZENNY SENT YOU THE PEARLS.**

**03)** (You two are preparing a special dinner for Zenny.)

Read it silently: How does Zenny like vinegar?

Say it aloud: **ZENNY DISLIKES VINEGAR.**

**04)** (Your coworker has heard that someone loves him/her.)

Read it silently: Who loves me?

Say it aloud: **ZENNY LOVES YOU.**

**05)** (At the coffee break, talking about who Ana loves.)

Read it silently: Who does Ana love?

Say it aloud: **ANA LOVES ZENNY.**

**06)** (After the visit of the government to the company.)

Read it silently: Who did the government talk to?

Say it aloud: **THE GOVERNMENT TALKED TO ZENNY.**

**07)** (Your coworker has to buy a ticket for someone from Pennsylvania back to Canada – your current location.)

Read it silently: Who traveled to Pennsylvania?

Say it aloud: **ZENNY TRAVELED TO PENNSYLVANIA.**

**08)** (Talking about the project of a ‘world’ your coworker created for a computer game.)

Read it silently: How does the boss like my ‘world’?

Say it aloud: **THE BOSS LOVES YOUR WORLD.**

**09)** (Talking about the way Zenny feels about your coworker.)

Read it silently: How does Zenny feel about me?

Say it aloud: **ZENNY LOVES YOU.**

**10)** (Talking about a T-shirt that you placed in the drier.)

Read it silently: How much did the T-shirt shrink?

Say it aloud: **THE T-SHIRT SHRANK A LOT.**

**Appendix K: Task 2 – I don't know the answer, do you! (Listener's version)**



Remember that you work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English. For this task, you will consider that you have previously asked your coworker some questions. By the time you asked the questions, your coworker didn't know the answers. However, after a little time, your coworker came out with the answers. Your job is to listen to the answers and figure out what question he or she is answering.

To illustrate, let us consider the following answer.

**Adele will be in Canada IN DECEMBER.**

The words in capital letters indicate that IN DECEMBER is emphasized. This emphasis means that IN DECEMBER is the most important information in the sentence and is the information you were looking for. In your sheet, you will have the questions you have previously asked your coworker, as illustrated below.

- (A) Who will be in Canada in December?**
- (B) Where will Adele be in December?**
- (C) When will Adele be in Canada?**

Then you should check alternative (C) for the question being answered, as the answer given is highlighting neither where (in Canada) nor who (Adele), but rather when (in December).

**START**

- 01)**  
(A) Who loves me?  
(B) How does Zenny feel about me?  
(C) Who does Zenny love?
- 02)**  
(A) Who sent the pearls to me?  
(B) What did Zenny send me?  
(C) Who did Zenny send the pearls to?
- 03)**  
(A) Who dislikes vinegar?  
(B) How does Zenny like vinegar?  
(C) What does Zenny dislike?
- 04)**  
(A) Who loves me?  
(B) How does Zenny feel about me?  
(C) Who does Zenny love?
- 05)**  
(A) Who loves Zenny?  
(B) How does Ana feel about Zenny?  
(C) Who does Ana love?
- 06)**  
(A) Who talked to Zenny?  
(B) What did the government do to Zenny?  
(C) Who did the government talk to?
- 07)**  
(A) Who traveled to Pennsylvania?  
(B) What about Zenny and Pennsylvania?  
(C) Where did Zenny travel to?
- 08)**  
(A) Who loves my 'world'?  
(B) How does the boss like my 'world'?  
(C) What did the boss love?
- 09)**  
(A) Who loves me?  
(B) How does Zenny feel about me?  
(C) Who does Zenny love?
- 10)**  
(A) What shrank a lot?  
(B) What happened to the T-shirt?  
(C) How much did the T-shirt shrink?

### Appendix L: Task 3 – You're wrong... (Speaker's version)



Remember that you work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English. For this task, you will consider different sentences, by reading them silently. Each sentence establishes a context that requests you to correct a piece of information given by your coworker. You will read aloud the utterance that corrects the information.

To illustrate, let us see an example below.

(Talking about your coworker John.)

Read it silently: John bought a car.

Say it aloud: **NO, JOHN BOUGHT A BIKE.**

When you say it, you emphasize BIKE, because that's the information your coworker is mistaken about (a bike, not a car).

**START**

**01** (Talking about who had the hair curled at the beauty salon yesterday.)

Read it silently: Ana had her hair curled.

Say it aloud: **NO, MARIA HAD IT CURLED.**

**02** (Talking about the colors that are appetite stimulators.)

Read it silently: Blue is an appetite stimulator.

Say it aloud: **NO, RED IS AN APPETITE STIMULATOR.**

**03** (Choosing the best rhythm to be played at the surprise party for Zenny.)

Read it silently: The rhythm suits Ana.

Say it aloud: **NO, THE RHYTHM SUITS ZENNY.**

**04** (Deciding on a pet to give to Zenny.)

Read it silently: Zenny loves cats.

Say it aloud: **NO, ZENNY LOVES DOGS.**

**05** (Talking about a proposal Zenny received (work in Brazil).)

Read it silently: Unfortunately Zenny disagreed.

Say it aloud: **NO, UNFORTUNATELY ZENNY AGREED.**

**06** (Talking about an unfortunate email sent to the boss.)

Read it silently: You sent the email.

Say it aloud: **NO, ANA SENT THE EMAIL.**

**07** (Deciding on a pet to give to Ana.)

Read it silently: Ana hates dogs.

Say it aloud: **NO, ANA LOVES DOGS.**

**08** (At coffee break, talking about general things.)

Read it silently: Pearls shine in vinegar.

Say it aloud: **NO, PEARLS MELT IN VINEGAR.**

**09** (Talking about who had the hair curled at the beauty salon yesterday.)

Read it silently: Ana had her hair curled.

Say it aloud: **NO, MARIA HAD IT CURLED.**

**10** (Talking about colors that suppress optimism.)

Read it silently: Yellow suppresses optimism.

Say it aloud: **NO, YELLOW STIMULATES OPTIMISM.**

### Appendix M: Task 3 – You're wrong... (Listener's version)



Remember that you work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English. You are a very-updated person. During your coffee break, you are having a conversation with your coworker about different subjects. Even though you are that updated, some information that you provide seems to be mistaken and your coworker corrects you. For this task, you will hear the information from your coworker and identify the piece being corrected.

To illustrate, consider that your coworker say:

**No, John bought a BIKE.**

The words in capital letters indicate that BIKE is emphasized. This emphasis means that BIKE is the information replacing the mistaken information provided by you. In your sheet, you will have possible corrections, as illustrated below:

- (A) John did, not Peter.
- (B) John bought it, he did not sell it.
- (C) A bike, not a car.

Then you should check alternative (C), as the speaker is correcting neither the agent (John) nor the action (bought), but rather the item purchased (a bike).

**START**



- 01)  
(A) Maria did, not Ana.  
(B) Maria had it curled, she didn't curl it herself.  
(C) Curled, not straightened.
- 02)  
(A) Red is, not blue.  
(B) Appetite, not optimism.  
(C) A stimulator, not a suppressor.
- 03)  
(A) The rhythm, not the place.  
(B) The rhythm suits Zenny, it doesn't annoy him.  
(C) It suits Zenny, not Ana.
- 04)  
(A) Zenny does, not Maria.  
(B) Zenny loves dogs; he doesn't hate them.  
(C) Dogs, not cats.
- 05)  
(A) Unfortunately, not 'fortunately'.  
(B) Zenny did, not Ana.  
(C) Zenny agreed; he didn't disagree.
- 06)  
(A) Ana did, not me.  
(B) Ana sent it, she didn't delete it.  
(C) The email, not the letter.
- 07)  
(A) Ana does, not me.  
(B) Ana loves dogs; she doesn't hate them.  
(C) Dogs, not cats.
- 08)  
(A) Pearls, not diamonds.  
(B) Pearls melt; they don't get shinier.  
(C) In vinegar, not in alcohol.
- 09)  
(A) Maria did, not Ana.  
(B) Maria had it curled, she didn't curl it herself.  
(C) Curled, not straightened.
- 10)  
(A) Yellow does, not green.  
(B) Yellow stimulates; it doesn't suppress.  
(C) Optimism, not negativism.

**Appendix N: Task 4 – I know what you’re saying next! (Speaker’s version)**



Remember that you work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English. For this task, you will consider different questions, by reading them silently. Each question represents a whole question that you want to ask. Note, however, that your

coworker likes finishing his/her interlocutors’ questions and will not allow you to finish yours. To illustrate, let us consider the example below.

(Talking about Maria’s trip.)

Read it silently:      Is Maria traveling to Canada tomorrow or to England?

Ask it aloud:        **IS MARIA TRAVELING TO CANADA TOMORROW ...?**

The part you read silently is the question you would ask, but you will not. The part you read aloud is the question you actually ask, as if your coworker were interrupting you.

When you ask it, you emphasize **CANADA**, because your second half would be another destination, namely **ENGLAND**.

**START**

**01)** (Deciding on the menu of the surprise party for Zenny.)

Read it silently: Does Zenny like shrimp or does he dislike it?

Ask it aloud: **DOES ZENNY LIKE SHRIMP ...?**

**02)** (Talking about the sequence of the names in the company's guest list for the surprise party for Zenny.)

Read it silently: Does this sequence represent the hierarchy or does it neglect it?

Ask it aloud: **DOES THIS SEQUENCE REPRESENT THE HIERARCHY ...?**

**03)** (Deciding on whom to invite to go parachuting on the weekend.)

Read it silently: Is Zenny afraid of heights or is it Ana?

Ask it aloud: **IS ZENNY AFRAID OF HEIGHTS ...?**

**04)** (Talking about the results of a survey carried out at the company.)

Read it silently: Do men lie a lot or is it women who do?

Ask it aloud: **DO MEN LIE A LOT ...?**

**05)** (Talking about the atmosphere at work when Zenny is around.)

Read it silently: Is the atmosphere around Zenny good or bad?

Ask it aloud: **IS THE ATMOSPHERE AROUND ZENNY GOOD ...?**

**06)** (Talking about a meeting with the boss.)

Read it silently: Does the meeting start at ten or does it end?

Ask it aloud: **DOES THE MEETING START AT TEN ...?**

**07)** (Talking about the exact moment when Zenny fell in love with Ana.)

Read it silently: Did Ana twirl her hair or her dress?

Read it aloud: **DID ANA TWIRL HER HAIR ...?**

**08)** (Before calling a taxi for a coworker.)

Read it silently: Is Zenny leaving immediately or is it Ana?

Read it aloud: **IS ZENNY LEAVING IMMEDIATELY ...?**

**09)** (Deciding on the menu of the surprise party for Zenny.)

Read it silently: Does Zenny like shrimp or does he dislike it?

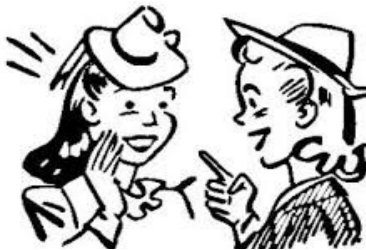
Ask it aloud: **DOES ZENNY LIKE SHRIMP ...?**

**10)** (Talking about the surprise party for Zenny.)

Read it silently: Does the party start at ten or at eight?

Read it aloud: **DOES THE PARTY START AT TEN ...?**

**Appendix O: Task 4 – I know what you're saying next! (Listener's version)**



Remember that you work in an International Company settled in Canada. Therefore, you have coworkers from different nationalities and your contact language is English. You are very sensitive and smart. You are fast at getting what people are saying and you love to finish your interlocutors' lines. For this task,

your coworker will ask you many questions, but you will hear only the first half of the questions as if you were interrupting him/her. Your job is to figure out what the second half would be, based on the way your coworker says it. To illustrate, let us consider that your coworker ask the question below.

**Is Maria traveling to CANADA tomorrow ...?**

The words in capital letters indicate that CANADA is emphasized. Thus, your coworker knows that Maria is traveling tomorrow but he/she is in doubt about the destination. In your sheet, you will have three possible second halves, as illustrated below.

- (A) or is it John?**
- (B) or to England?**
- (C) or the day after tomorrow?**

Then you should check alternative (B), as the speaker is in doubt about neither the agent (Maria or John) nor the day (tomorrow or the day after tomorrow), but rather the destination (Canada or England).

**START**

**01)**

- (A) or is it Anna?
- (B) or does he dislike it?
- (C) or salmon?

**02)**

- (A) or is it the other sequence?
- (B) or does it neglect the hierarchy?
- (C) or the beloved ones?

**03)**

- (A) or is it Anna?
- (B) or is he crazy about it?
- (C) or of dogs?

**04)**

- (A) or is it women who do?
- (B) or do they tell the truth?
- (C) or just a little?

**05)**

- (A) or is it the people around Zenny who are good?
- (B) or around Anna?
- (C) or bad?

**06)**

- (A) or is it the party?
- (B) or does it end?
- (C) or at eight?

**07)**

- (A) or was it Maria?
- (B) or did she tie her hair?
- (C) or her dress?

**08)**

- (A) or is it Ana?
- (B) or is he staying?
- (C) or later?

**09)**

- (A) or is it Anna?
- (B) or does he dislike it?
- (C) or salmon?

**10)**

- (A) or is it the meeting?
- (B) or does it end?
- (C) or at eight?

## Appendix P: Word Familiarity Test



In a scale from 1 to 9, check how well you know these words. 1 (one) means you **don't know** the word **at all** and 9 (nine) means that you **know** the word **very well** and you have even used it in your interactions in English.

Your name: \_\_\_\_\_

Word	Scale
01) Afraid	(1) (2) (3) (4) (5) (6) (7) (8) (9)
02) Agreed	(1) (2) (3) (4) (5) (6) (7) (8) (9)
03) Appetite	(1) (2) (3) (4) (5) (6) (7) (8) (9)
04) Atmosphere	(1) (2) (3) (4) (5) (6) (7) (8) (9)
05) Bike	(1) (2) (3) (4) (5) (6) (7) (8) (9)
06) Boss	(1) (2) (3) (4) (5) (6) (7) (8) (9)
07) Bought	(1) (2) (3) (4) (5) (6) (7) (8) (9)
08) Curled	(1) (2) (3) (4) (5) (6) (7) (8) (9)
09) December	(1) (2) (3) (4) (5) (6) (7) (8) (9)
10) Dislikes	(1) (2) (3) (4) (5) (6) (7) (8) (9)
11) Dogs	(1) (2) (3) (4) (5) (6) (7) (8) (9)
12) Good	(1) (2) (3) (4) (5) (6) (7) (8) (9)
13) Government	(1) (2) (3) (4) (5) (6) (7) (8) (9)
14) Had	(1) (2) (3) (4) (5) (6) (7) (8) (9)
15) Hair	(1) (2) (3) (4) (5) (6) (7) (8) (9)
16) Hates	(1) (2) (3) (4) (5) (6) (7) (8) (9)

17) Heights	(1) (2) (3) (4) (5) (6) (7) (8) (9)
18) Hierarchy	(1) (2) (3) (4) (5) (6) (7) (8) (9)
19) Immediately	(1) (2) (3) (4) (5) (6) (7) (8) (9)
20) Leaving	(1) (2) (3) (4) (5) (6) (7) (8) (9)
21) Lie	(1) (2) (3) (4) (5) (6) (7) (8) (9)
22) Like	(1) (2) (3) (4) (5) (6) (7) (8) (9)
23) Loves	(1) (2) (3) (4) (5) (6) (7) (8) (9)
24) Meeting	(1) (2) (3) (4) (5) (6) (7) (8) (9)
25) Melt	(1) (2) (3) (4) (5) (6) (7) (8) (9)
26) Men	(1) (2) (3) (4) (5) (6) (7) (8) (9)
27) Optimism	(1) (2) (3) (4) (5) (6) (7) (8) (9)
28) Party	(1) (2) (3) (4) (5) (6) (7) (8) (9)
29) Pearls	(1) (2) (3) (4) (5) (6) (7) (8) (9)
30) Pennsylvania	(1) (2) (3) (4) (5) (6) (7) (8) (9)
31) Red	(1) (2) (3) (4) (5) (6) (7) (8) (9)
32) Rhythm	(1) (2) (3) (4) (5) (6) (7) (8) (9)
33) Sequence	(1) (2) (3) (4) (5) (6) (7) (8) (9)
34) Shrank	(1) (2) (3) (4) (5) (6) (7) (8) (9)
35) Shrimp	(1) (2) (3) (4) (5) (6) (7) (8) (9)
36) Start	(1) (2) (3) (4) (5) (6) (7) (8) (9)
37) Stimulator	(1) (2) (3) (4) (5) (6) (7) (8) (9)
38) Suits	(1) (2) (3) (4) (5) (6) (7) (8) (9)
39) Talked	(1) (2) (3) (4) (5) (6) (7) (8) (9)
40) Ten	(1) (2) (3) (4) (5) (6) (7) (8) (9)
41) Tomorrow	(1) (2) (3) (4) (5) (6) (7) (8) (9)

42) Traveled	(1) (2) (3) (4) (5) (6) (7) (8) (9)
43) Traveling	(1) (2) (3) (4) (5) (6) (7) (8) (9)
44) T-shirt	(1) (2) (3) (4) (5) (6) (7) (8) (9)
45) Twirl	(1) (2) (3) (4) (5) (6) (7) (8) (9)
46) Unfortunately	(1) (2) (3) (4) (5) (6) (7) (8) (9)
47) Vinegar	(1) (2) (3) (4) (5) (6) (7) (8) (9)
48) World	(1) (2) (3) (4) (5) (6) (7) (8) (9)
49) Yellow	(1) (2) (3) (4) (5) (6) (7) (8) (9)



## Appendix Q: Pronunciation Self-Evaluation Test



In a scale from 1 to 9, check how well you pronounce these words. 1 (one) means that you are sure you **don't know** how to pronounce the word **at all** and 9 (nine) means that you are sure that **you know** how to pronounce the word **very well**.

Your name: \_\_\_\_\_

Word	Scale
01) Afraid	(1) (2) (3) (4) (5) (6) (7) (8) (9)
02) Agreed	(1) (2) (3) (4) (5) (6) (7) (8) (9)
03) Appetite	(1) (2) (3) (4) (5) (6) (7) (8) (9)
04) Atmosphere	(1) (2) (3) (4) (5) (6) (7) (8) (9)
05) Bike	(1) (2) (3) (4) (5) (6) (7) (8) (9)
06) Boss	(1) (2) (3) (4) (5) (6) (7) (8) (9)
07) Bought	(1) (2) (3) (4) (5) (6) (7) (8) (9)
08) Curled	(1) (2) (3) (4) (5) (6) (7) (8) (9)
09) December	(1) (2) (3) (4) (5) (6) (7) (8) (9)
10) Dislikes	(1) (2) (3) (4) (5) (6) (7) (8) (9)
11) Dogs	(1) (2) (3) (4) (5) (6) (7) (8) (9)
12) Good	(1) (2) (3) (4) (5) (6) (7) (8) (9)

13)	Government	(1) (2) (3) (4) (5) (6) (7) (8) (9)
14)	Had	(1) (2) (3) (4) (5) (6) (7) (8) (9)
15)	Hair	(1) (2) (3) (4) (5) (6) (7) (8) (9)
16)	Hates	(1) (2) (3) (4) (5) (6) (7) (8) (9)
17)	Heights	(1) (2) (3) (4) (5) (6) (7) (8) (9)
18)	Hierarchy	(1) (2) (3) (4) (5) (6) (7) (8) (9)
19)	Immediately	(1) (2) (3) (4) (5) (6) (7) (8) (9)
20)	Leaving	(1) (2) (3) (4) (5) (6) (7) (8) (9)
21)	Lie	(1) (2) (3) (4) (5) (6) (7) (8) (9)
22)	Like	(1) (2) (3) (4) (5) (6) (7) (8) (9)
23)	Loves	(1) (2) (3) (4) (5) (6) (7) (8) (9)
24)	Meeting	(1) (2) (3) (4) (5) (6) (7) (8) (9)
25)	Melt	(1) (2) (3) (4) (5) (6) (7) (8) (9)
26)	Men	(1) (2) (3) (4) (5) (6) (7) (8) (9)
27)	Optimism	(1) (2) (3) (4) (5) (6) (7) (8) (9)
28)	Party	(1) (2) (3) (4) (5) (6) (7) (8) (9)
29)	Pearls	(1) (2) (3) (4) (5) (6) (7) (8) (9)
30)	Pennsylvania	(1) (2) (3) (4) (5) (6) (7) (8) (9)
31)	Red	(1) (2) (3) (4) (5) (6) (7) (8) (9)
32)	Rhythm	(1) (2) (3) (4) (5) (6) (7) (8) (9)
33)	Sequence	(1) (2) (3) (4) (5) (6) (7) (8) (9)
34)	Shrank	(1) (2) (3) (4) (5) (6) (7) (8) (9)
35)	Shrimp	(1) (2) (3) (4) (5) (6) (7) (8) (9)
36)	Start	(1) (2) (3) (4) (5) (6) (7) (8) (9)
37)	Stimulator	(1) (2) (3) (4) (5) (6) (7) (8) (9)

38) Suits	(1) (2) (3) (4) (5) (6) (7) (8) (9)
39) Talked	(1) (2) (3) (4) (5) (6) (7) (8) (9)
40) Ten	(1) (2) (3) (4) (5) (6) (7) (8) (9)
41) Tomorrow	(1) (2) (3) (4) (5) (6) (7) (8) (9)
42) Traveled	(1) (2) (3) (4) (5) (6) (7) (8) (9)
43) Traveling	(1) (2) (3) (4) (5) (6) (7) (8) (9)
44) T-shirt	(1) (2) (3) (4) (5) (6) (7) (8) (9)
45) Twirl	(1) (2) (3) (4) (5) (6) (7) (8) (9)
46) Unfortunately	(1) (2) (3) (4) (5) (6) (7) (8) (9)
47) Vinegar	(1) (2) (3) (4) (5) (6) (7) (8) (9)
48) World	(1) (2) (3) (4) (5) (6) (7) (8) (9)
49) Yellow	(1) (2) (3) (4) (5) (6) (7) (8) (9)

## Appendix R: Praat pitch tracking issues

Some pitch contours were not accurately tracked by Praat, such as in *Ana loves ZENNY* (Task 2) produced by B07 (Figure R1). Observe that on *Zenny* there are two tracing for the pitch contour. One starting down at 90.6 st and the other a little higher at 92.8 st.

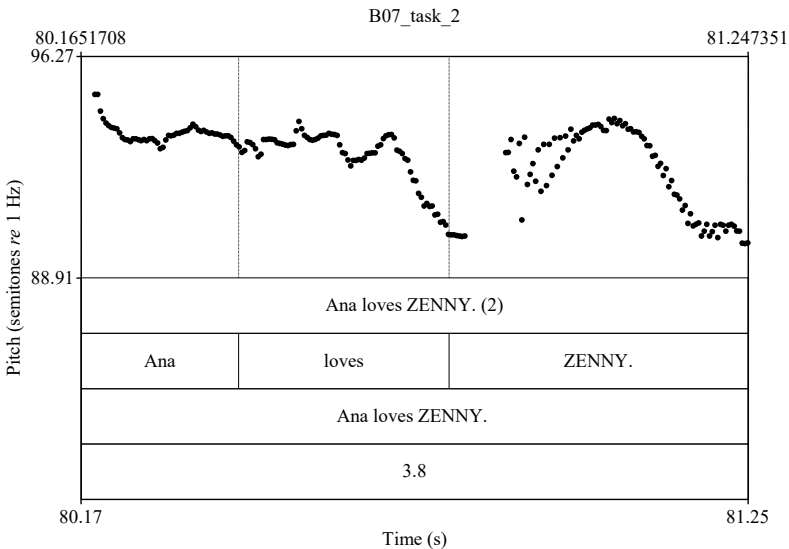


Figure R1: Production of *Ana loves ZENNY* (Task 2) by speaker B07 – An example of different pitch-track tracing by Praat.

Another illustration is in *Zenny DISLIKES vinegar* (Task 2, Item 3), produced by B09. Praat could not track the F0 properly probably due to creaky voice on *dislikes*. In these cases, the researcher had to listen to the production and discern the falling or rising direction of the pitch based on the researcher's own perception of the tones.

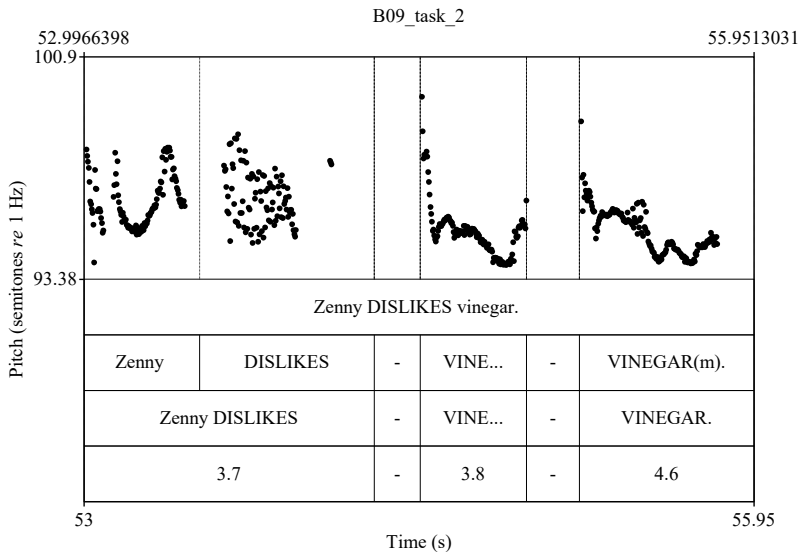


Figure R2: Production of *Zenny DISLIKES vinegar* (Task 2, Item 3), produced by speaker B09 – An example of different pitch-track tracing by Praat.

There were cases in which it seemed that the last emphasis was just that of marking the boundary tone. However, in order to be consistent with the criteria for data analysis, I had to consider and measurement the pitch change because it was preceded by a pause longer than 200 ms, which signaled the beginning of a new thought group. One example is provided in Figure R3.

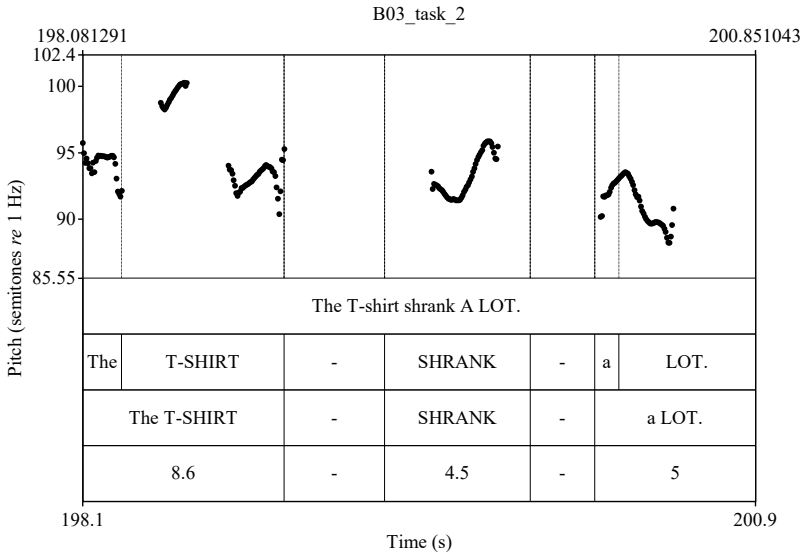


Figure R3: Production of *The T-shirt shrank A LOT* (Task 2, Item 10) by B03 – An example of confusing final pitch change.

In this production, there were two pauses. One between *T-shirt* and *shrank* and another between *shrank* and *a lot*. Their durations are 414 and 265ms, respectively. Observe that in *The T-SHIRT*, there seems to be low boundary tone as if there were nothing else to say, while in *shrank*, there is rising intonation, which in this case seems to signal that something else is coming (listing effect). The last intonational phrase *a LOT* has a low boundary tone, closing the sentence. If there were no pauses between *shrank* and *a lot*, it would be considered an only thought group with only one NS, as the pitch contour would start at *shrank* and finish in *a lot*.

## Appendix S: Dismissed Cases

Some productions were not considered in the interpretation of the speaker's intent for two reasons: (1) the speaker said more than the speaker was expected to say; (2) the speaker did not understand the discursive context and thus did not understand what portion of the utterance should hold the NS. Figures S1 and S2 illustrates these two cases.

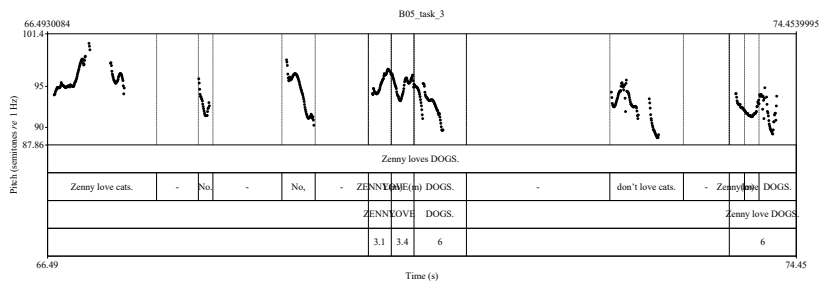


Figure S1: *Zenny loves DOGS* (Task 3, Item 04) by B05 – An example of a dismissed case (1).

The context for the utterance in Figure S1 was a conversation to decide what pet to give Zenny. The speaker had to correct that Zenny loves dogs rather than cats. During the production of the utterance, B05 got confused and replaced the word *dogs* with the word *cats* and then corrected herself by saying more than she should (“No, Zenny love DOGS. Don’t love cats. Zenny love DOGS”). As a result, it is impossible to measure if the correct answer was due to the correct production of NS (the last time she said it – look at the rightmost part of the figure) or fruit of her explanation. Therefore, it was not considered in the interpretation analysis.

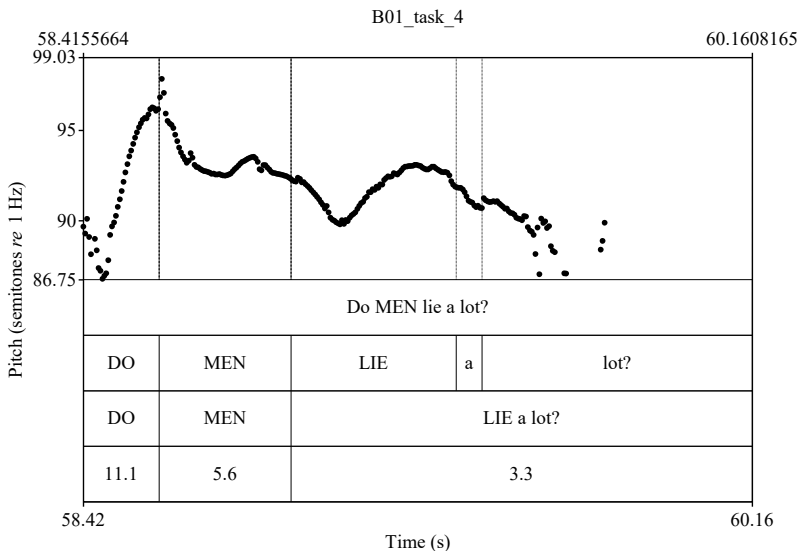


Figure S2: *Do MEN lie a lot?* (Task 4, Item 04) by B1 – An example of a dismissed case (2).

Figure S2 gives an example of a production that was dismissed because the speaker did not understand which word should be nuclear stressed. The context for this question was a conversation about the results of a survey carried out at the company. The speaker had to ask the question as if he were in doubt about the results for who lies more: men or women. Therefore, he should nuclear stress *men*. According to her highlights in her production sheet, done after the completion of the interaction tasks, she understood that she should nuclear stress the word *lie*, meaning that she was in doubt if men lie or tell the truth. Figure S2 shows the production of three pitch changes: *do* (11.1 st), *men* (5.6 st), and *lie* (3.3 st). Observe that the word meant to hold the NS had the smallest pitch change and that the word that should be nuclear stressed has no pitch change at all. Although a rise was expected on *lot* as speakers in this dataset tended to rise at the end to signal questions, *lot* has a low boundary tone (L%). Despite the unintended meaning in this production, the listener (F01) perceived the NS in *a lot* and interpreted that B01 was in doubt about the amount of lies told by men: a lot or just a little. *Lot* was the word with more duration, which may have contributed to the perception of it as more emphatic. F01 reported that



the pronunciation of segments (strong *t* at the end, probably produced because of duration) hindered the perception of NS: “But maybe it’s because of the word *lot* and the strong T at the end that she says...” (Interview with F01, 00:32:56 – 00:33:01). This production together with the other 12 that fall into this category were dismissed.

## Appendix T: Complexity of words and NS placement – more illustrations and some qualitative analysis

NS was more often placed as expected in utterances with non-challenging words than in utterances with challenging words and the difference, according to a chi-square test for group independence, was found to be statistically significant. Table T1 displays a more detailed account of the first-time productions with regard to pronunciation and some prosodic variables, such as pauses and intonation, prosodic factors that are also important to the perception of thought groups and thus of NS. In the column 'status', the aspects are displayed. They were labeled as follows: (1) *Exp. Neutral*: expected productions with no apparent pronunciation issues; (2) *Exp. + misp.*: expected productions with pronunciation issues; (3) *Unexp. neutral*: unexpected productions with no apparent pronunciation issues; (4) *Unexp. Pron.*: unexpected productions with apparent pronunciation issues; (5) *Unexp. pauses*: unexpected productions with pauses; (6) *Unexp. pron. + pauses*: unexpected productions with apparent pronunciation issues and pauses; (7) *Unexp. Rep. + pron.*: unexpected productions with pronunciation issues and repetition; (8) *Unexp. rep. + pron. + pauses*: unexpected productions with pronunciation issues, repetition, and pauses; (9) *Unexp. SL/H*: unexpected productions with unexpected syllable lengthening or rising intonation; (10) *Unexp. SL + pause*: unexpected productions with unexpected syllable lengthening or rising intonation and pauses; (11) *Unexp. SL + pron.*: unexpected productions with syllable lengthening or rising intonation and unexpected pronunciation; (12) *Unexp. + SL + pause + pron.*: unexpected productions with syllable lengthening or rising intonation, unexpected pronunciation, and pauses; (13) *Unexp. + SL + pause + pron. + rep.*: unexpected productions with syllable lengthening or rising intonation, unexpected pronunciation, pauses, and repetition. The next paragraphs are dedicated to explaining and exemplifying each aspect. Some figures were previously used in Section 4.1.3. They are repeated here as well as their discussion in order to facilitate the reading and comprehension.

Table T1  
*Pronunciation details – first-time production*

Status	Not		Short		Long		Total	
	challenging		Challenging		Challenging			
	N	%	N	%	N	%	N	%
(1)Exp. neutral	18	14.3	6	4.8	1	.8	25	6.6
(2)Exp. + misp.	15	11.9	9	7.1	14	11.1	39	10.0
(3)Unexp. neutral	30	23.8	19	15.1	7	5.6	56	14.8
(4)Unexp. Pron.	17	13.5	21	16.7	13	10.3	50	13.5
(5)Unexp. pauses	4	3.2	3	2.4	0	.0	7	1.9
(6)Unexp. Pron. + pauses	8	6.3	10	7.9	17	13.5	35	9.3
(7)Unexp. Rep. + pron.	0	.0	2	1.6	4	3.2	6	1.6
(8)Unexp. Rep. + pron. + pauses	5	4.0	7	5.6	8	6.3	20	5.3
(9)Unexp. SL/H	9	7.1	5	4.0	10	7.9	24	6.3
(10)Unexp. SL + pause	1	.8	4	3.2	1	.8	6	1.6
(11)Unexp. SL + pron.	15	11.9	17	13.5	28	22.2	60	15.9
(12)Unexp. SL + pause + pron.	1	.8	17	13.5	13	10.3	31	8.2
(13)Unexp. + SL + pause + pron. + rep.	3	2.4	6	4.8	10	7.9	19	5.0
Total	126	100.0	126	100.0	126	100.0	378	100.0

The expected productions with no pronunciation issues (Aspect 1) were more frequent in utterances with simple words (14.3%, N = 18), followed by utterances with short challenging words (4.8%, N = 6) and then by long challenging words (.8%, N = 1) which had a sole aspect. This signals that for these speakers placing NS was challenging most of

the times, but it was less complicated when challenging words were absent. Figure T1 (originally Figure 13, Section 4.1.1), repeated and renamed here for easy retrieval, illustrates an expected production of an utterance with non-challenging words.

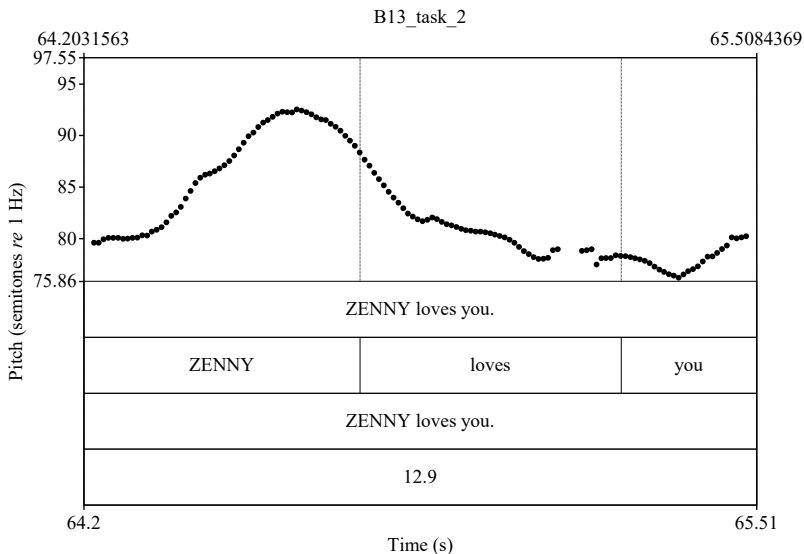


Figure T1: *ZENNY loves you* (Task 2, Item 04) by B13 – An example of an expected NS placement in utterances with non-challenging words.

Figure T1 shows the production of *ZENNY loves you* by B13. Note that none of the words in the utterance should pose a challenge in terms of pronunciation for B13. The only word that could have had an expected pronunciation is *loves*, with vowel change or vowel epenthesis, but those were not the case. The easy words in terms of pronunciation here may have helped B13 to nuclear-stress the expected word, namely, *Zenny*, which was, in turn, perceived by F13, who grabbed the speaker's intent in this specific case. Figures T2 and T3 (originally Figures 19 and 20, Section 4.1.3) below illustrate productions with no pronunciation issues in utterances with short and long words, respectively.

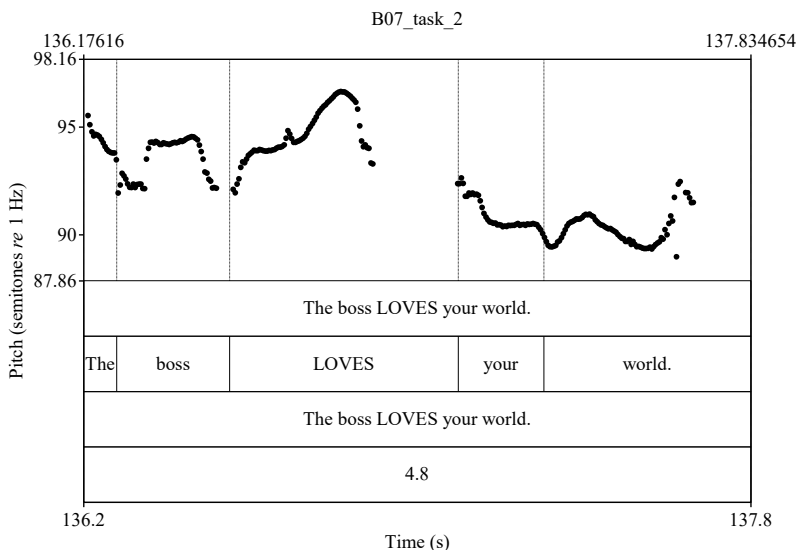


Figure T2: *The boss LOVES your world* (Task 2, Item 08) by B07 – An example of an expected NS production with no pronunciation issues in utterances with short-challenging words.

The context for the production in Figure T2 was that of business. People were talking about the project of a ‘world’ created by a coworker for a computer game and the speaker was expected to say the utterance as a response to the context question *How does the boss like my ‘world’?* and thus, the word *loves* should be nuclear stressed. The production was delivered with no pauses or pronunciation problems. Although some time was spent in saying *world*, *loves* had the pitch change of 4.8 semitones and more duration. The word *world* had many pronunciations in the dataset: [wɜrd], [wɜrdʒ], [wɜ:d], [ɔrd], and [wɔud]. Even though B07 rated this word a 3 (three), which means she was not sure if her pronunciation for this word was intelligible enough, the word did not pose any obstacles: it was pronounced as expected and the NS was also assigned as expected.

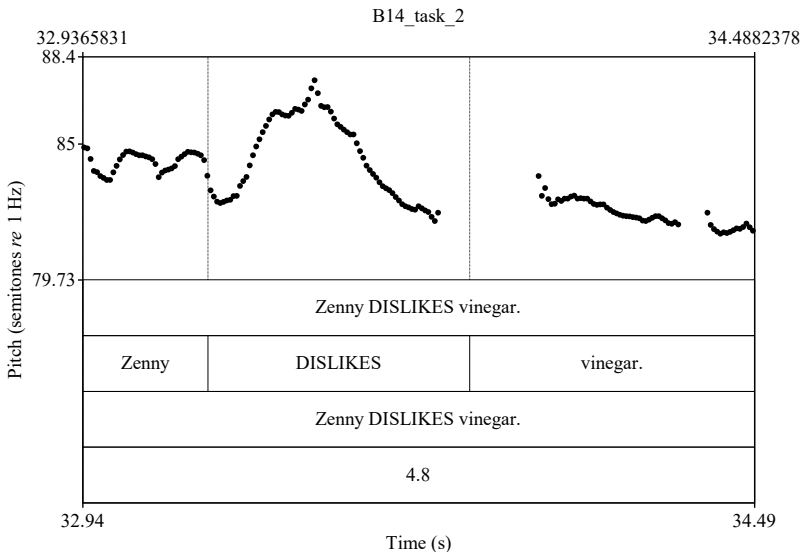


Figure T3: *Zenny DISLIKES vinegar* (Task 2, Item 03) by B14 – An example of an expected NS production with no pronunciation issues in utterances with short-challenging words.

The context for the production in Figure T3 was the preparation of special dinner for Zenny, the character created for data collection. B14 had to say the utterance in reply to the context question *How does Zenny like vinegar?* and thus the word *dislikes* should be nuclear stressed. The challenging word here was *vinegar*, which was hardly ever pronounced as expected along the dataset, being often pronounced as [ˈvaɪnɪgər] or [ˈvaɪnəgər], and less often as [vɪˈnɛgər], [ˈvɪnɛgər], and [ˈvɪnɪgər]. B14 rated her pronunciation of *vinegar* a 9 (nine), which means she was very confident that the way she pronounced the word was very intelligible. It may have helped her with the accurate assignment of NS.

Table T1 also displays the number of times the expected productions had some pronunciation issues (Aspect 2). This time, the frequencies were not so distant from one another, being a little more frequent for utterances with simple words (11.9%, N = 15), followed by utterances with long words (11.1%, N = 14) and then by utterances with short words (7.1%, N = 9). The difference between them may lie on the nature of the phonological processes involved in the unexpected

pronunciations, rendering one (non-challenging words) more intelligible than the others (short and long-challenging words). Such productions are depicted in Figures T4, T5, and T6.

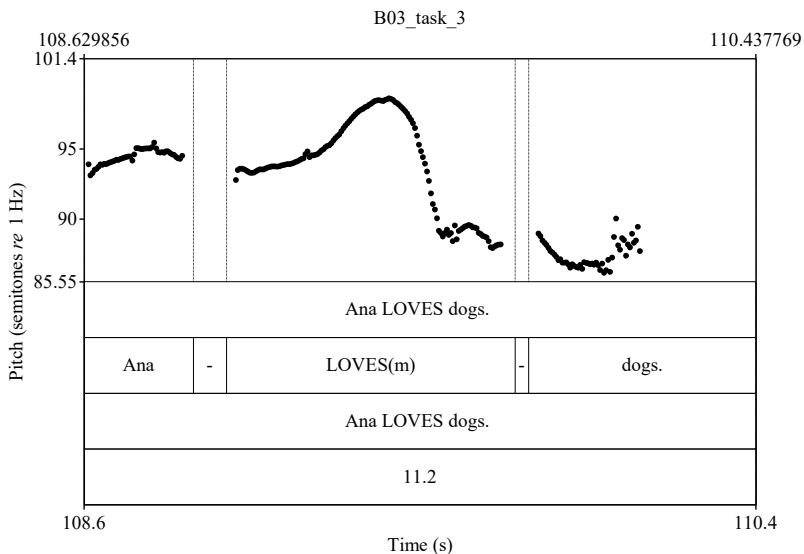


Figure T4: *No, Ana LOVES dogs* (Task 2, Item 07) by B03 – An example of an expected NS production with pronunciation issues in utterances with non-challenging words.

In this production, B03 nuclear stressed the word *loves* to imply that Ana does not hate dogs. Her pitch change was that of 11.2 st and was completely clear. Her pronunciation of *loves* came out as [lavz], which is a very common pronunciation among Brazilians. *Loves* was also pronounced in the dataset as [lavz], [lɔvz], [ˈlɔvɪz], [lav], [lavz], which all seemed to be intelligible enough. In this specific production, the unexpected pronunciation did no harm to communication as F03 got the expected message.

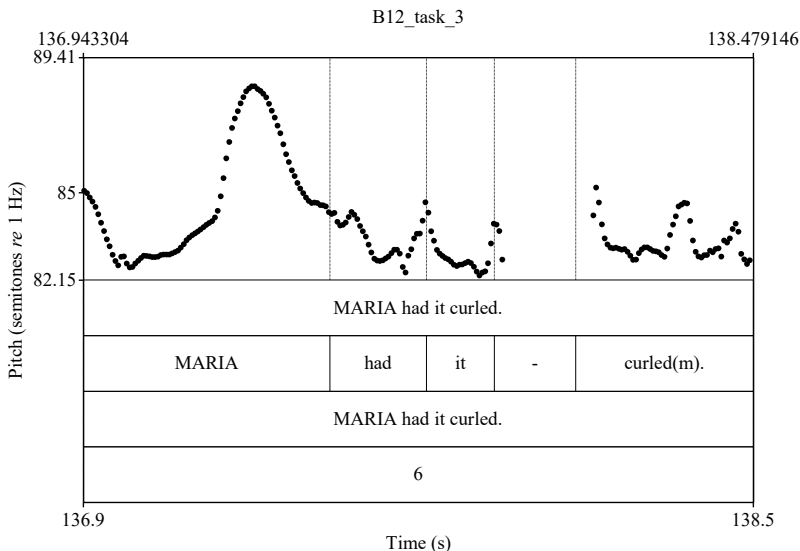


Figure T5: *No, MARIA had it curled* (Task 3, Item 09) by B12 – An example of an expected NS production with pronunciation issues in utterances with short-challenging words.

According to the context set for the production in Figure T5, B12 was supposed to nuclear-stress *Maria* in order to correct the mistaken information that Ana had her hair curled. In terms of pitch change and NS placement, the utterance was produced as expected with a six-semitone-pitch change on *Maria*. There was a short pause right before *curled*, shorter than 200 ms and was not considered. The word *curled* was pronounced [<sup>1</sup>kɜrlɪd], which distracted the listener who asked B12 to repeat. In the repetition, her production was alike the first with regard to both pitch change (6.4 st) and pronunciation ([<sup>1</sup>kɜrlɪd]), but F12 could concentrate on the NS this time and get the expected message. [<sup>1</sup>kɜrlɪd] was the most frequent pronunciation for *curled* in the dataset. Other pronunciations for this word were [kɪld] and [<sup>1</sup>kurɪr], which sound to be a way less intelligible. B12 found that her pronunciation of the word could be not intelligible enough, and that was so.



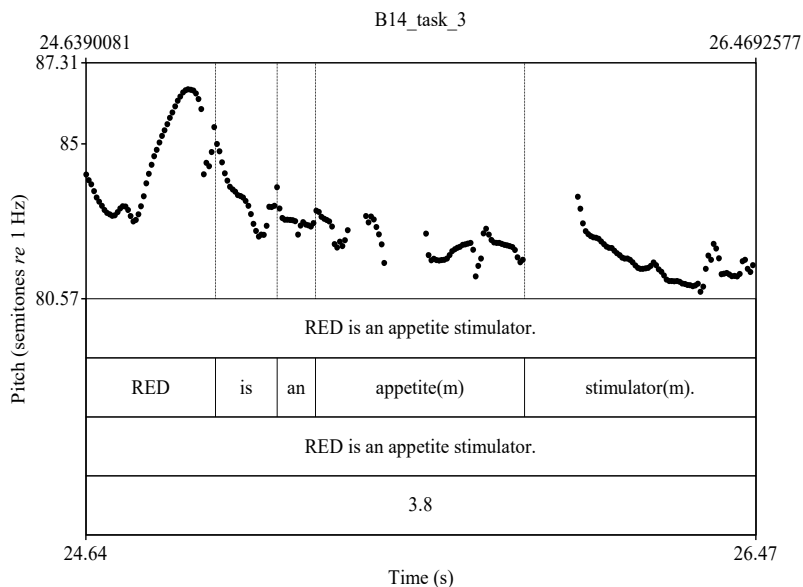


Figure T6: *No, RED is an appetite stimulator* (Task 3, Item 02) by B14 – An example of an expected NS production with pronunciation issues in utterances with long-challenging words.

The context for the production in Figure T6 was a conversation about world knowledge in which a coworker mentions that blue is an appetite stimulator. The speaker had to correct the information by saying that red rather than blue was an appetite stimulator, by putting the NS on *red*. The challenging word here in terms of pronunciation was meant to be *appetite*. However, *stimulator* also showed to be a hard word for these Brazilians to pronounce. B14 rated her pronunciation of these words as totally intelligible and that probably helped her with the correct placement of NS, as she was totally confident about pronouncing them: No obstacles were present. Her pronunciations of these words were [ˈæpɪˌtɑɪrə] and [ˈsɪmjʊleɪrə], respectively, not intelligible enough as her pronunciation changes the words severely, mainly the latter. These unexpected pronunciations did not hinder the expected perception of the NS on *red*<sup>79</sup>. Other pronunciations of these words were [ˈæpɪˌtɪt],

<sup>79</sup> Many times the listeners reported that they could perceive the emphasis but they were unable to understand the rest of the utterance. This just give us food for thought. Recall that Jenkins

[ˈʌpɪ,tɪt], [ˈʌmpə,taiʃɪ], [ˈæ::pɪtɪt], [ʌˈpeɪtɪt], and [ˌʌpeˈtɪt] for *appetite*, and [ˈsɪmuleɪrə], [ˈsɪmjuleɪtər], [ˈsɪmjʊ,leɪə], [ˈstʃɪmjuleɪtʃər], [ˈɪstɪmjʊˈleɪrə], [ɪsˈtɪmjʊ,leɪtə], [ˈɪstɪmʊˈleɪrər], and [ˌsɪmʊˈleɪrər] for *stimulator*, with apparent different levels of intelligibility.

Some NS assignment were unexpected, but had no pronunciation issues (Table T1, Aspect 3). They were more frequent in utterances with non-challenging words (23.8%, N = 30), followed by utterances with short challenging words (15.1%, N = 19) and then by utterances with long challenging words (5.6%, N = 7). It is a coherent result, since non-challenging words are easier to pronounce. This fractal information gives away what the general results in Table 13 (Section 4.1) showed: Placing NS for these Brazilians was a difficult and challenging task. During data collection, when receiving the explanation of the tasks, most of them asked the researcher how they were supposed to emphasize the information, a piece of evidence that they were not aware of how this phenomenon would take place. Figures T7, T8, and T9 illustrate these unexpected productions.

---

(2000) gives an example of an IL-English speaker interacting with other members of the international community and asking for a blue pen, by saying “Do you have a blue *vun*?”, *vun* meaning *one*. Jenkins argues that this speaker’s interlocutors could not make sense of what was being said because *vun* was nuclear stressed. If the speaker had placed the NS on blue (which contrasts with another color), she argues, the listeners would have understood that he wanted a blue pen. However, my point here is that it might not be the case. If the listeners in my study did not have the options, would they have been able to grab the speaker’s intent? The case discussed and illustrated in Figure T6 points to a no. The listener understood only the word being corrected (*red*), so he got that the speaker’s intent was to correct the color, but he was not able to understand the message at all. Who can guarantee that those interactants in Jenkins’s study would understand right way what was being said? It is possible that as they did not have any options to check, they would try to understand the whole question and getting the meaning of *vun* was important. Future studies could disentangle this issue by having utterances being heard by two groups: with and without the options.

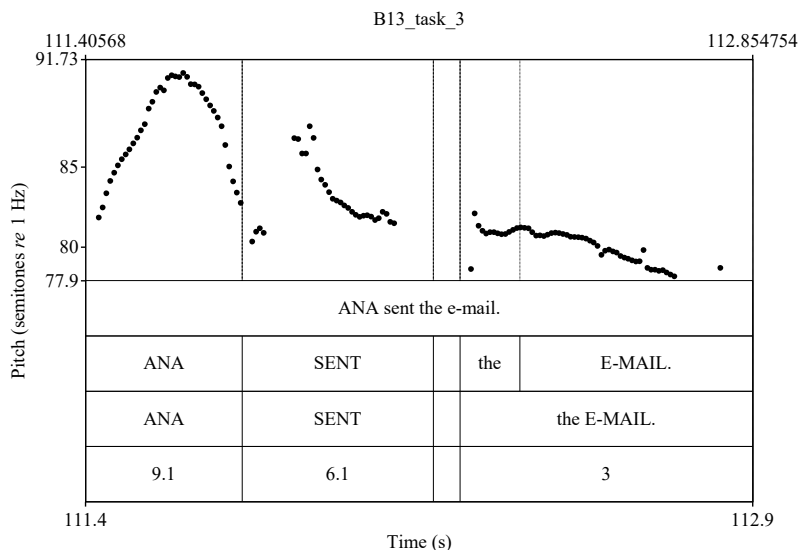


Figure T7: *ANA sent the e-mail* (Task 2, Item 06) by B13 – An example of an unexpected NS production with no pronunciation issues in utterances with non-challenging words.

The context for this utterance (Figure T7) is a chat about an unfortunate e-mail sent to the boss. The speaker has to correct that Ana rather than the speaker himself sent the e-mail, by placing NS on *Ana*. B13 had a wide pitch range and his NS were usually loud and clear. However, because of his inflections, many other words were salient. Observe that the NS on *Ana* has a pitch change of 9.1 st, but he was unable to totally deemphasize the remainder of the sentence, with two other pitch changes on *sent* and *e-mail* (6.1 and 3st, respectively), the latter probably caused by the final boundary tone. This production was interpreted as expected by the listener.

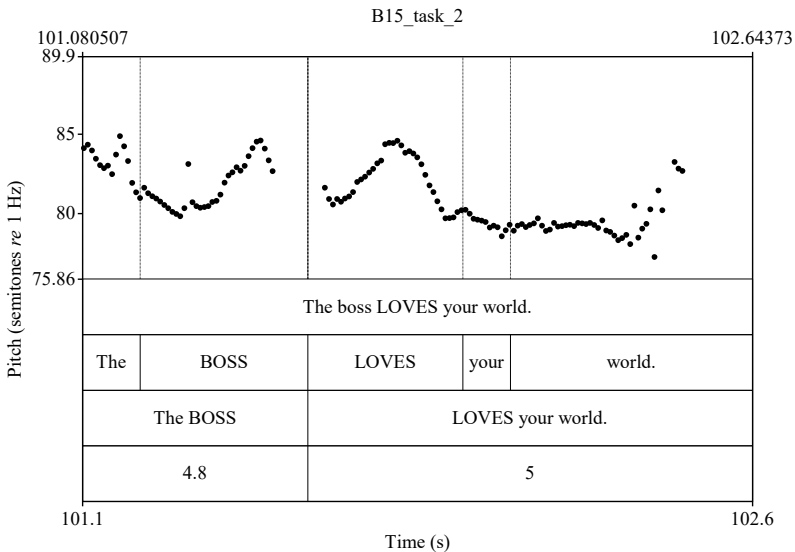


Figure T8: *The boss LOVES your world* (Task 2, Item 08) by B15 – An example of an unexpected NS production with no pronunciation issues in utterances with short-challenging words.

In the utterance in Figure T8, B15 had to nuclear stress the constituent *loves* in order to convey the information of how the boss liked the listener's world for the new computer game. This production was mispronunciation free. However, two similar pitch changes could be tracked on *boss* and *loves* (4.8 and 5 st, respectively). The short challenging word here was *world*, which was rated as intelligibly enough in terms of pronunciation and was pronounced as expected. The two pitch changes however could have been triggered by the care and attention to the pronunciation ahead, as *world* showed to be the word with greater duration. F15 asked B15 to repeat because the two NSs were not distinct for her and she got to notice the expected NS with the repetition. Nevertheless, when B15 repeated the utterance, the pronunciation of *world* came out as [ɔrd]. This indicates that caring about pronunciation interfered with his expected production. Even though the repetition was also an unexpected production, *loves* had a pitch change a little higher than did *boss* (0.8 st) and more duration than any word in the utterance, which made it sound undoubtedly more prominent.

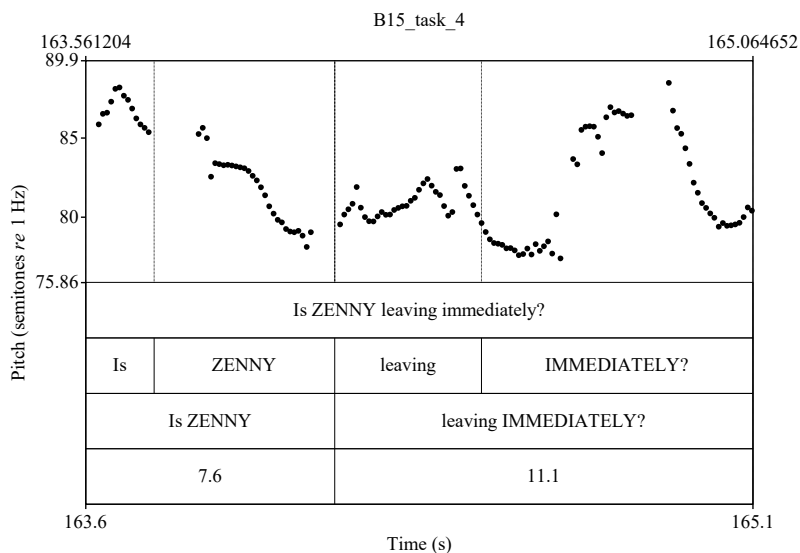


Figure T9: *Is ZENNY leaving immediately?* (Task 4, Item 08) by B15 – An example of an unexpected NS production with no pronunciation issues in utterances with long-challenging words.

B15 had as a context the duty of calling a taxi for a coworker, but he did not know if the coworker was Zenny or Ana, as both would be needing a taxi, but only one would leave at that moment. The entire question was *Is ZENNY leaving immediately or is it ANA?* and B15 had to ask only the first half of the question as if F15 interrupted him, preventing him from completing the question. Figure T9 shows that B15 placed NS on *Zenny* with a pitch change of 7.6 semitones. However, it was a falling pitch change instead of an expected rising, as it is a question. In order to make it sound like a question, he had another pitch change of 11.1 semitones in the final portion of the utterance, doing the circumflex movement, similar to that of the BP pattern for neutral polar questions (Morales, 2008). His challenge here was intonation rather than pronunciation. *Immediately* was the long-challenging word and B15 thought his pronunciation of this word was intelligible enough (rated as 8). Despite the fact that the greatest pitch change, duration and eyebrow movement were on *immediately*, as depicted from the videos, the listener was able to get the expected interpretation, but was unable to explain why.

In Aspect 4, Table T1, one can see the results for productions with both unexpected NS assignments and unexpected pronunciations. They represent 16.7% (N = 21) of utterances with short-challenging words, followed by 13.5% (N = 17) of utterances with non-challenging words and then by utterances with long-challenging words (10.3%, N = 13). Figures T10, T11, and T12 (originally Figures 21, 22 and 23, Section 4.1.3) illustrate these productions.

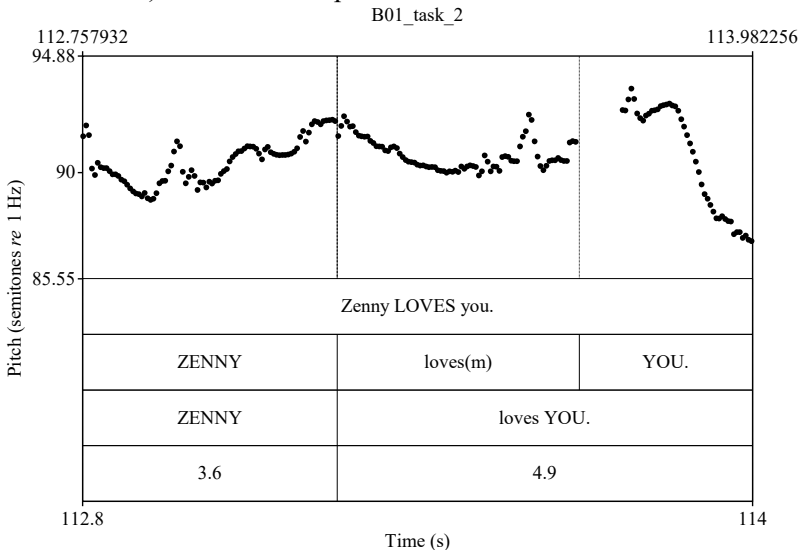


Figure T10: *Zenny LOVES you* (Task 2, Item 04) by B01 – An example of a production with both unexpected NS placement and unexpected pronunciation in utterances with non-challenging words.

The expected NS location for the utterance in Figure T10 is on *loves*, in order to imply how Zenny feels about the listener. However, two sights of pitch change are found. *Zenny* has some prominence (a 3.6-st-pitch change) and *you* has the greatest pitch change (4.9 st). Emphasizing destressed words such as *you* is a common practice among Brazilians, as duly noted by Baptista (2001). The only word that should take pitch change here was *loves*, but it was destressed and pronounced as [l'ɔvɪz]. This exemplifies the difficult some participants had in placing NS in the middle of the sentence, which goes in line with Carpes's (2014) results. F01 got an unexpected interpretation for this

production: that the question being answered was ‘*Who does Zenny love?*’.

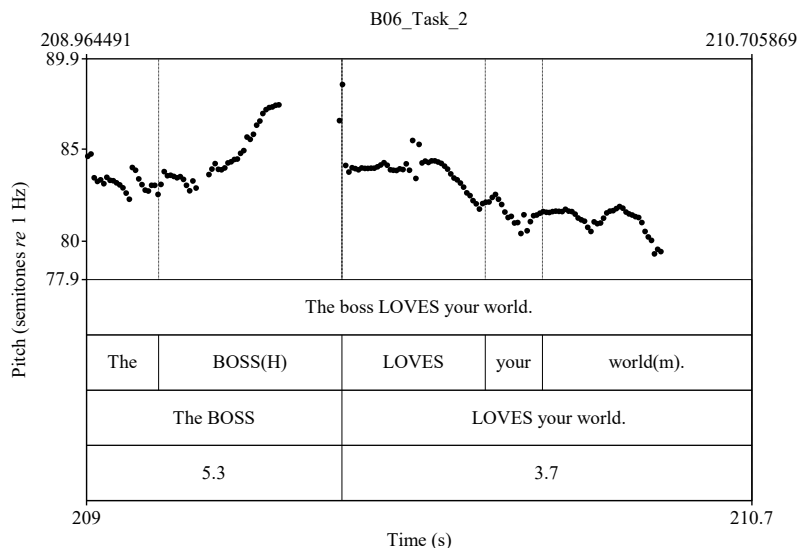


Figure T11: *The boss LOVES your world* (Task 2, Item 08) by B06 – An example of a production with both unexpected NS placement and unexpected pronunciation in utterances with short-challenging words.

The expected location for NS in the utterance in Figure T11 is also *loves*, this time to imply how the boss feels about the world created by the listener to a computer game. Observe that there is some rising intonation on *boss*, with a 5.3-semitone-pitch change. This gives the impression that there is more to be said but, at the same time, misguides the listener with regard to the NS. *Loves*, the expected portion for NS assignment, also has a pitch change, but smaller than the previous one: 3.7 semitones. *World* is mispronounced here in a way that the researcher was unable to phonetically represent it. Unexpected pronunciations in the dataset were [wɜrd], [wɜrdʒ], [wɜ:d], [ɔrd], and [wɔud]. Listening to this production once did not enable F06 to interpret the utterance. Consequently, she asked B06 to repeat. It is my impression that when he repeated, B06 was more concerned with correcting his pronunciation than to placing the NS correctly, as he also nuclear stressed *world*, pronounced [wɜr] in the repetition, with a 4.4-semitone-pitch change. F06 then understood that they were talking about the object of the boss’s

love rather than about how the boss felt in relation to the *world* created for the computer game.

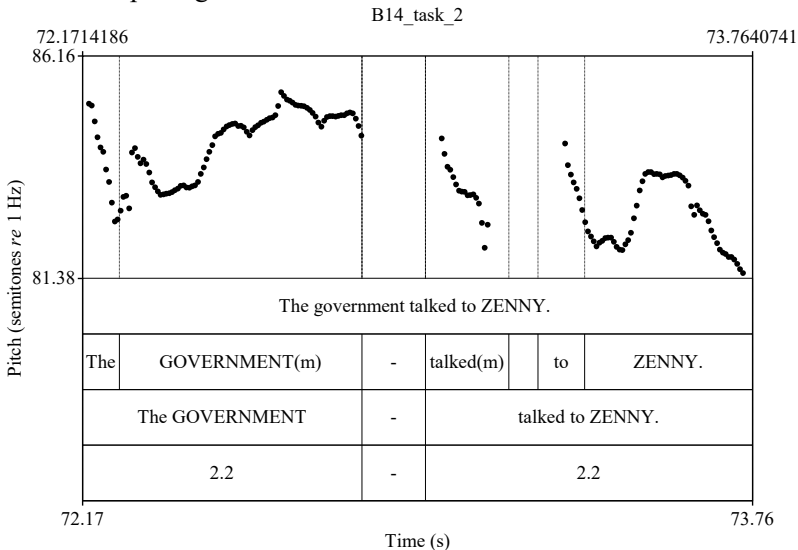


Figure T12: *The government talked to ZENNY* (Task 2, Item 06) by B14 – An example of a production with both unexpected NS placement and unexpected pronunciation in utterances with long-challenging words.

The context for this utterance was a visit of the government to the company. It was known that the government talked to someone specific, but it was not clear to whom. So, *Zenny* should be nuclear stressed. This specific production is particularly interesting as no portion of the utterance reaches a change up to three semitones. The two visible contours on *government* and *Zenny* are exactly the same: 2.2 semitones. The word *government* was produced as [ˈgʌvər̩ment] and *talked* as [tɔk]. Other pronunciations found in the dataset were [ˈgʊvər̩ment], [ˈgʊvɜr̩ment] [ˈgʊvər̩nt], [ˈgʌvər̩ment], [ˈgʊvər̩nəment], [ˈgarvər̩ment] and [tɔk], [ˈtɔkəd], [ˈtɔkəd], [tɔkd], and [ˈtɔkɪd]. B14 thought that her pronunciation of *government* would probably be unintelligible, but that of *talked* would be completely intelligible. It is difficult to say what exactly happened here, but it is my impression that, being a quiet person, instead of getting loud on the word she was not sure about, B14



may have felt intimidated and used a lower tone for the whole sentence. Despite this, F14 could curiously get the expected message<sup>80</sup>.

There were instances in which unexpected productions had no pronunciation issues, but they had unexpected pauses (Table T1, Aspect 5). They represent only 1.9% of the dataset (N = 7): 3.2% (N = 4) of the utterances with non-challenging words and 2.4% (N = 3) of the utterances with short-challenging words. Figure T13 provides an example of this production.

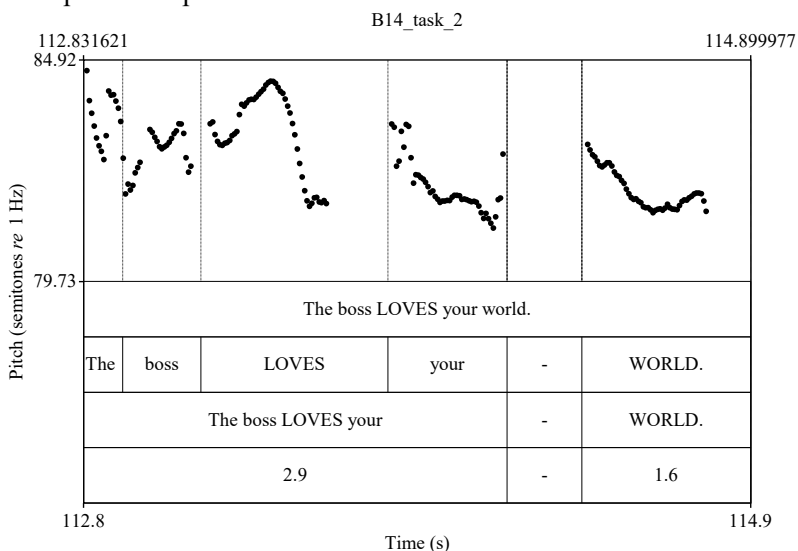


Figure T13: *The boss LOVES your world* (Task 2, Item 08) by B14 – An example of an unexpected production with unexpected pauses.

NS should fall on *loves* in the utterance in Figure T13. The challenging word for B14 here is *world*, whose pronunciation she rated as intelligible enough. Indeed, hers was an expected pronunciation for the word. Although her pronunciation was adequate, it caused her to pause<sup>81</sup> (231ms) before the word, breaking the utterance into two thought groups: *The boss loves\* your* and *world\**. However, observe

<sup>80</sup> In this case, there was no way of comparing duration of the words, as they had quite different length, and no extra cues were used during the delivery of this utterance.

<sup>81</sup> According to Fernandes-Svartman (2008), this pause before a nuclear-stressed word is a characteristic of topic. However, analyzing whether pauses may be connected to topicalization goes beyond the scope and design of the present study.

that neither *loves* nor *world* had a clear NS as the pitch changes were not greater than 2.9 semitones. This production corroborates my impression about B14's strategy of using a more plain tone when facing challenging words to pronounce.

Unexpected NS productions with both unexpected pronunciation and unexpected pauses (Table T1, Aspect 6) were more frequent than with pauses alone (Table T1, Aspect 5). They represent 9.3% of the utterances in the dataset (N = 35). They were more frequent in utterances with long-challenging words (13.5%, N = 17), followed by those with short-challenging words (7.9%, N = 10) and then by those with non-challenging words (6.3%, N = 8). Figure T14 gives an example of this type of production.

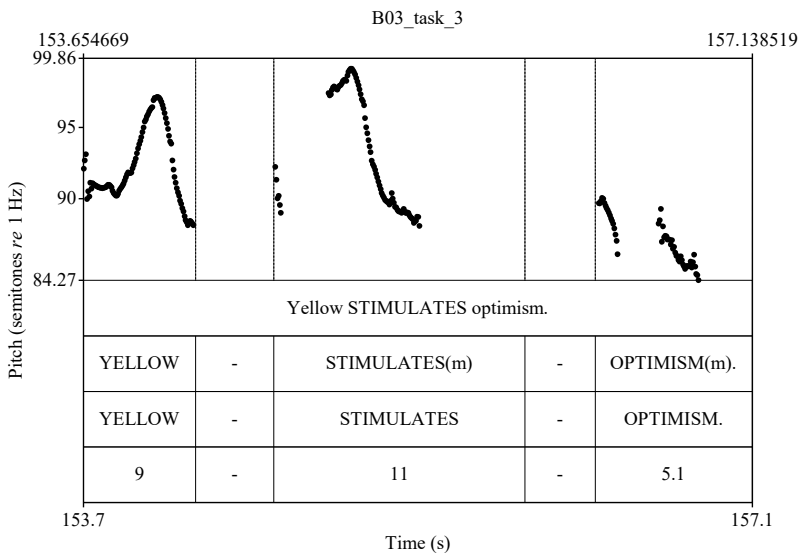


Figure T14: *Yellow STIMULATES optimism* (Task 2, Item 10) by B03 – An example of an unexpected NS production with both unexpected pronunciation and unexpected pauses.

The context for the production in Figure T14 was an informal chat on world knowledge. The speaker was told that yellow suppresses optimism, and she has to correct the information by saying that yellow stimulates rather than suppresses optimism, by placing NS on *stimulates*. The figure shows that three distinct pitch movements were produced: *yellow* (9-semitone-pitch change), *stimulates* (11-semitone-

pitch change) and *optimism* (5.1-semitone-pitch change). There were two pauses: (1) between *yellow* and *stimulates* and (2) between *stimulates* and *optimism*, of 407 ms and 368 ms, respectively. Pausing before hard words to pronounce showed to be a tendency in the dataset, be the pause accompanied or not by syllable lengthening<sup>82</sup>, mispronunciation, and/or repetition. *Stimulates* was not a word that the researcher expected to be difficult to pronounce but that showed to be a challenge. It was pronounced as [ɪs'tɪmʊ,leɪtɪs] by B03 (vowel epenthesis + vowel change). Other pronunciations found in the dataset were: ['sɪmjʊleɪtɪs], ['ɪstɪmjʊ,leɪt], [,stɪmʊ'leɪtɪs], [ɪs'tʃɪmʊ,leɪtɪs], ['ɪstɪmʊ,leɪtɪs], ['sɪmʊ,leɪtɪs], and ['ɪsɪmʊ,leɪtɪs]. Her pronunciation for the expected challenging word was ['ɑptɪ,mɪst]. She rated her pronunciation of this word as intelligible enough (rating 5), but her pronunciation seems to be a little unintelligible in my point of view, given that it resembles the word *optimist* more than *optimism*. Other pronunciations in the dataset for this word were: ['ɑptɪ,mɪzɪm], ['ɑptɪmɪs], ['ɑptɪmɪz], ['ɑtʃɪ,mɪsmʊ], ['ɑptʃɪ,mɪsmʊ], ['ɑptɪmɪs], ['ɑptɪ,mɪs], [ɑpʊ'tʃɪmɪs], [ɑ:p'tʃɪsmʊ], [ɑ'petʃɪ,mɪs], ['ɑtʃɪ,mɪs], and ['ɑ,pɪns], being the latter apparently the least intelligible of them all. In fact, these two words seemed to be obstacles for B03 to place the NS as expected. F03 asked her to repeat. When she repeated, there was still a pause between *stimulates* and *optimism*, both were still mispronounced, but B03 was able to place a greater pitch change in *stimulates* (14.2 st) and she used extra cues, such as eyebrow movement and nodding when saying the expected word. As a result, F03 could get the intended message.

At some moments, repetition was combined with mispronunciation or with pauses and syllable lengthening or with the three. Repetition combined with mispronunciation (Table T1, Aspect 7) was produced only in utterances with short- (1.6%, N = 2) and long- (3.2%, N = 4) challenging words. Figure T15 illustrates the productions under this fashion.

---

<sup>82</sup> The production of unexpected pauses and unexpected syllable lengthening is a problem in terms of prosody processing, as both pauses and syllable lengthening are cues for thought group limits (Heusinger, 2007) and their unexpected use may break speech in nonsense units.

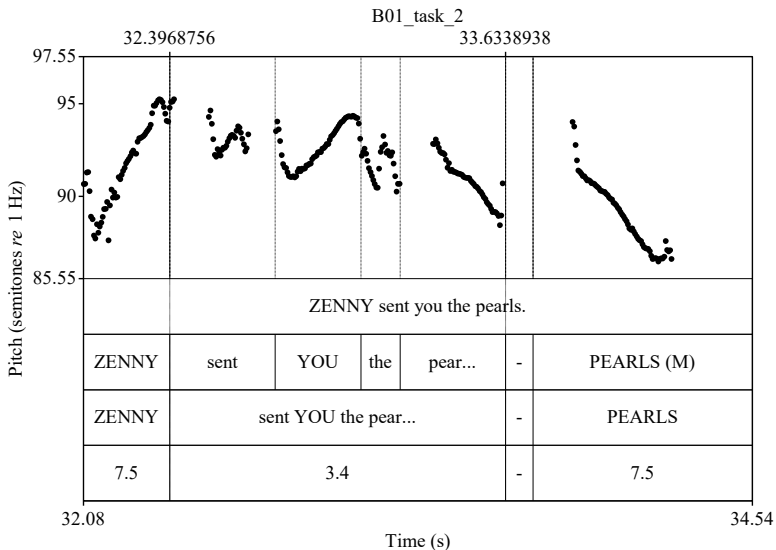


Figure T15: *Zenny sent you the pearls* (Task 2, Item 02) by B01 – An example of an unexpected NS production combined with unexpected repetition and unexpected pronunciation.

The context for the production in Figure T15 was a chat about the pearls the listener received from a secret admirer. B01 had to say the utterance in a response to the question *Who sent the pearls to me?* and thus nuclear stress *Zenny*. In this production, she places prominence on *Zenny* (7.5 st), *you* (3.4 st), and *pearls* (7.5 st). Note that *Zenny* and *pearls* hold the same pitch change. The challenging word here is *pearls*, whose pronunciation was rated as intelligible enough by B01. However, it caused her some trouble, as she started it as the word *pear* and it came out pronounced on the second try as [ˈpæris], the city. Other pronunciations for *pearls* in the dataset were: [pɪərs], [pɜrs], [pɪəʊs], [ˈpɪɜrs], [pɪəʊs], [peʊrs] and [pers], all quite unintelligible in my point of view. F01 reported that she could not perceive any distinction between the words in terms of prominence and she chose *pearls* as the emphasized information, meaning that the question being answered was *What did Zenny send me?* as if the listener did not know that she had received the pearls.

Repetitions associated with unexpected pronunciations and pauses were a little more recurrent (Table T1, Aspect 8). They were

more frequent in utterances with long-challenging words (6.3%, N = 8), followed by utterances with short-challenging words (5.6%, N = 7) and then by utterances with non-challenging words (4.0%, N = 5). Sometimes participants resorted to repetition in order to correct NS placement or pronunciation per se. Figure T16 displays an example of this kind of production.

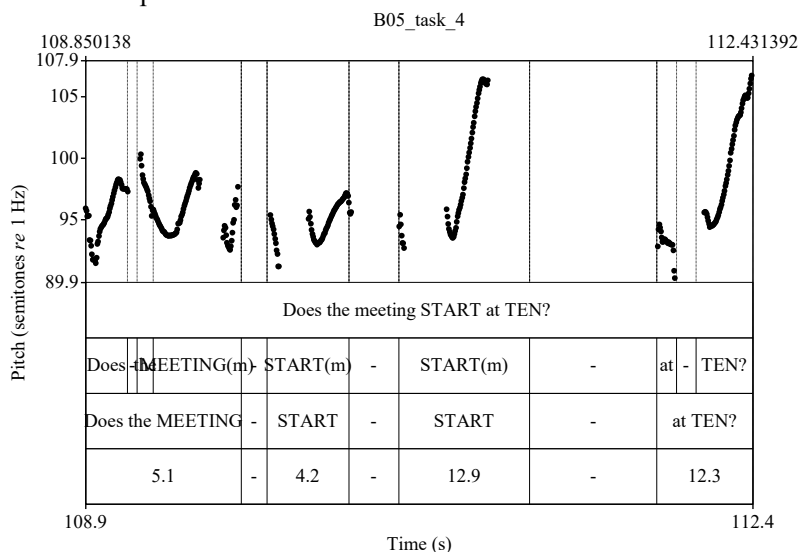


Figure T16: *Does the meeting START at ten?* (Task 4, Item 06) by B05 – An example of an unexpected NS production combined with unexpected repetition in an attempt to correct NS placement.

The context for the utterance in Figure T16 is the schedule of a meeting. The speaker is in doubt if the meeting starts or ends at ten. Therefore, B05 should nuclear stress the word *start*. Nevertheless, she said *start* (pronounced [ɪs'tɑrt]) under a small pitch change (4.2 st) and then corrected herself using a greater pitch change (12.9 st) with rising intonation. Instead of keeping her tone high, she lowered it to say *at ten* right after the 683-millisecond pause. *Ten* also had a rising intonation and a great pitch change of 12.3 semitones, but this time the rising intonation was not used to signal NS but the end of the question. F05, however, got that B05 was in doubt about the time being ten or eight rather than if the meeting would start or end at ten.

Another way of gaining time in order to process how to pronounce words occurred by means of syllable lengthening. This phenomenon would be expected at the NS word as more prominent words usually also have more duration (Calhoun, 2007; Shue et al., 2007). However, in the present dataset, it was also a resource used by speakers to gain more time or were used in words they were careful when pronouncing. Rising intonation was computed together with syllable lengthening as at many times they appeared in tandem with one another. These lengthening and/or rising intonation were sometimes used in tandem with pauses, mispronunciation, repetition or the combination of the three, which made speech difficult to process. The utterances with sole syllable lengthening/rising intonation were more frequent in utterances with long-challenging words (7.9%, N = 10), followed by those with non-challenging words (7.1%, N = 9) and then by those with short-challenging words (4.0, N = 5). Figure T17 illustrates productions with this process.

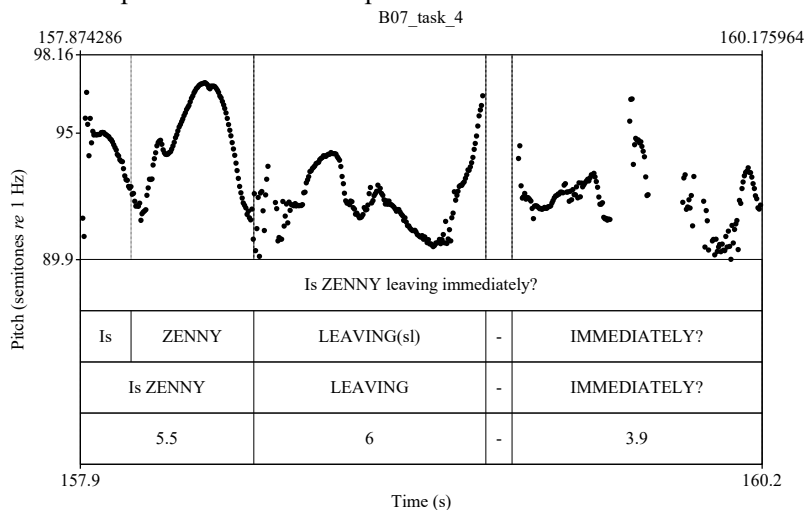


Figure T17: *Is ZENNY leaving immediately?* (Task 4, Item 21) by B7 – An example of an unexpected NS production combined with unexpected syllable lengthening and rising intonation.

In order to imply that the speaker was in doubt if it was Zenny or Ana who was leaving immediately, B07 had to nuclear stress *Zenny*. In her production, she put NS on *Zenny* (5.5 st), *leaving* (6 st), and *immediately* (3.9 st). *Zenny* had the pitch change probably because it

was the portion meant to be emphasized. The word *leaving* had some syllable lengthening probably caused either by the desire of asking a question or by the threat ahead: the hard long word *immediately*. She rated her pronunciation for the word as probably not intelligible; that rating explains the care. There were no mispronunciations in this production at the expenses of the production of three thought groups and three NSs. F07 asked her to repeat, but she still had *leaving* with syllable lengthening and, this time, *immediately* was pronounced [ɪ'mediətli]. Other pronunciations in the dataset were: [ɪ'miriəli], [ɪ'midiətli], [ɪ'medi,ɛtli], [ɪ'meditli], [ɪ'mi,dʌntli], [ɪ'midi'ɛtli], [ɪm'miri,ɛtli]; [ɪ'miri,ɛtli], [ɪ'midzi'ɛtli], [ɪ'midɪlɪ,tɪ], [ɪ'midi,ɛtli], [ɪ'medzi,ɛtli], and [ɪ'midi'ɛtli]. F07 did not manage to accommodate to these processes by B07 and got an unexpected interpretation of B07's intent.

Syllable lengthening combined with pauses (with expected pronunciations) were not frequent. They appeared in 3.2% (N = 4) of the utterances with short challenging words and .8% (N = 1) of utterances with either non- or long-challenging words. This type of production is illustrated in Figure T18.

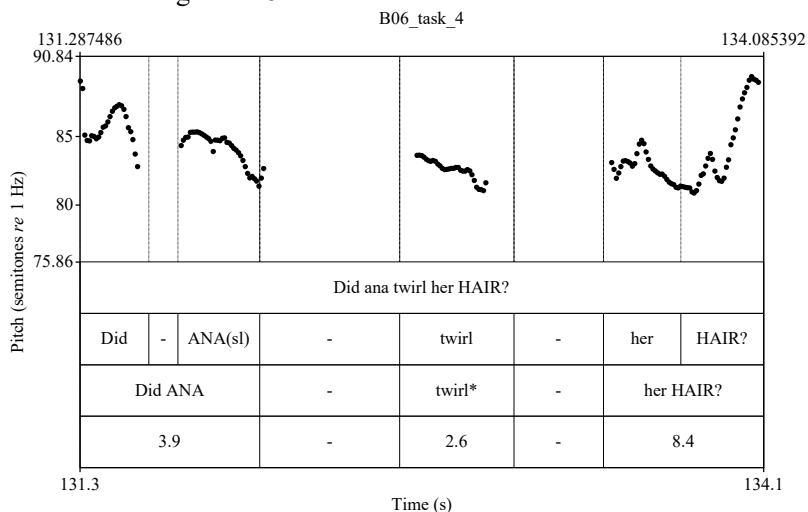


Figure T18: *Did ANA twirl her HAIR?* (Task 4, Item 07) by B06 – An example of an unexpected NS production combined with unexpected syllable lengthening and pauses (with expected pronunciation).

The utterance in Figure T18 had as a context a conversation about the exact moment when Zenny fell in love with Ana. The speaker was in doubt if at that moment Ana twirled her hair or her dress. Therefore, he should nuclear stress *hair* or *her hair* in his question. He produced two pauses: (1) between *Ana* (which has some syllable lengthening) and *twirl* (572 ms) and (2) between *twirl* and *her hair* (366 ms). These pauses broke the utterance into three thought groups: ‘Did *Ana*’ had a small pitch change (3.9 st), ‘*twirl*’ did not reach the three-semitone-pitch change (2.6 st), and ‘*her HAIR*’, which had the greatest pitch change in the utterance (8.4 st). Both syllable lengthening on *Ana* and pause after *Ana* may have been caused by the presence of *twirl*, a word that required B06’s attention. He rated his pronunciation for this word is probably unintelligible, but it came out with an expected pronunciation. Interestingly, even though *twirl* did not reach the pitch change for it to be perceived as holding the NS, F06 perceived the emphasis on that word, interpreting that B06 was in doubt if Ana had twirled or tied her hair. This perception may have been a result of perceived duration, given that with the caring pronunciation *twirl* had the longest duration in the utterance (467 ms), even longer than *Ana* that had syllable lengthening. The great pitch change (also with the question-rising intonation) and the extra cue (eyebrow movement) on *hair* were not enough to enable the listener to grab the NS on *hair*. Other pronunciations of *twirl* in the dataset were [tʃɜr], [ˈtʃɪrni], [θroʊl], [twerl], [tʃu], [twiʊ], [twel], [trɔʊ], [tʃuerl], [twiʊl], [twer], [tʃuˈaɪli], [tʃɪɾɔʊ], and [twiər], and they were carried out by most speakers in the present study (N = 13).

Utterances with both syllable lengthening and unexpected pronunciations (Table T1, Aspect 11) were far more frequent than those with syllable lengthening and expected pronunciations (Table T1, Aspect 10), especially for utterances with long- and short-challenging words. They were present in 22.2% (N = 28) of the former and in 13.5% (N = 17) of the latter, and in 11.9% (N = 15) of the utterances with non-challenging words. Figure T19 gives an example of these productions.



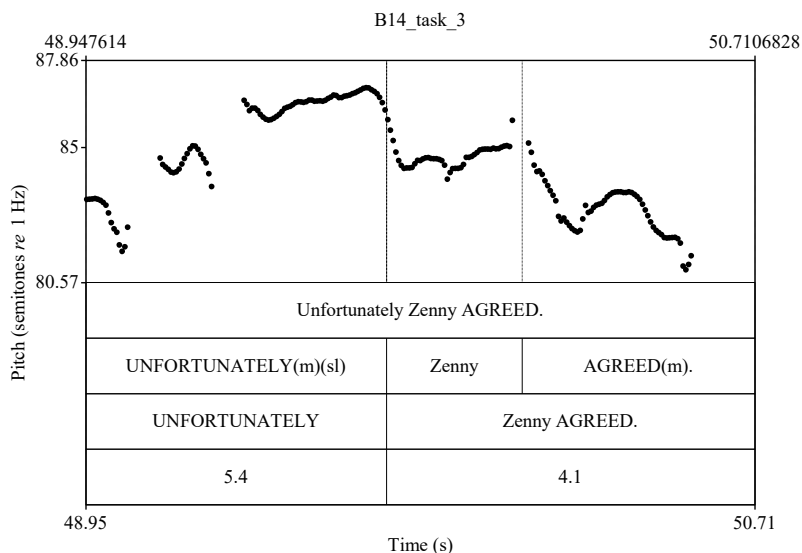


Figure T19: *Unfortunately Zenny AGREED* (Task 2, Item 05) by B14 – An example of an unexpected NS production combined with unexpected syllable lengthening/rising intonation and unexpected pronunciation.

The context for the utterance in Figure T19 is the fact that Zenny received a proposal to work in another country (Brazil). The speaker has to correct the information that Zenny disagreed with going to Brazil and has to place NS on *agreed*. B14 produced pitch changes on *unfortunately* (5.4 st), which also had both some mispronunciation ([ʌn'fɔrtʃənəli]) and syllable lengthening/rising intonation, and *agreed* (4.1st) pronounced [ɛ'grɪd]. The pronunciation of these two words was considered intelligible enough and totally intelligible by B14, respectively. F14 reported that he asked her to repeat because he could not perceive any distinct emphasis. His final answer, supposedly based on the repetition, was for NS on *agreed*. However, in the repetition that portion of the utterance did not reach the three-semitone-pitch change (2.6 st only). The two locations are not comparable in terms of duration, given that *unfortunately* was produced as [ʌn'fɔrtʃənəli] (5 syllables) and *agreed* with only two syllables. The video shows that there were no extra cues in this production. So, there is a chance that F14 could not perceive any distinct pitch change even after the repetition and was

guided by the feeling that the final portion of the utterance holds the NS and then he chose it.

Some productions had the combination of unexpected syllable lengthening, pauses and pronunciation (Table T1, Aspect 12). They represent 13.5% (N = 17) of the utterances with short-challenging words, 10.3% (N = 13) of the utterances with long-challenging words, and only .8% (N = 1) of the utterances with non-challenging words. This implies that the challenging words may have contributed to more noise in the productions, they were, as F13 mentioned, obstacles. This is corroborated by the results in Aspect 13 (Table T1), which displays the results for utterances that had the combination of unexpected syllable lengthening, pause, pronunciation, and repetition. They were more frequent in utterances with long-challenging words (7.9%, N = 10), followed by utterances with short-challenging words (4.8% of them, N = 6), and then by utterances with non-challenging words (2.4% of them, N = 3). Figures T20 and T21 illustrate the productions in Aspects 12 and 13 (Table T1), respectively.

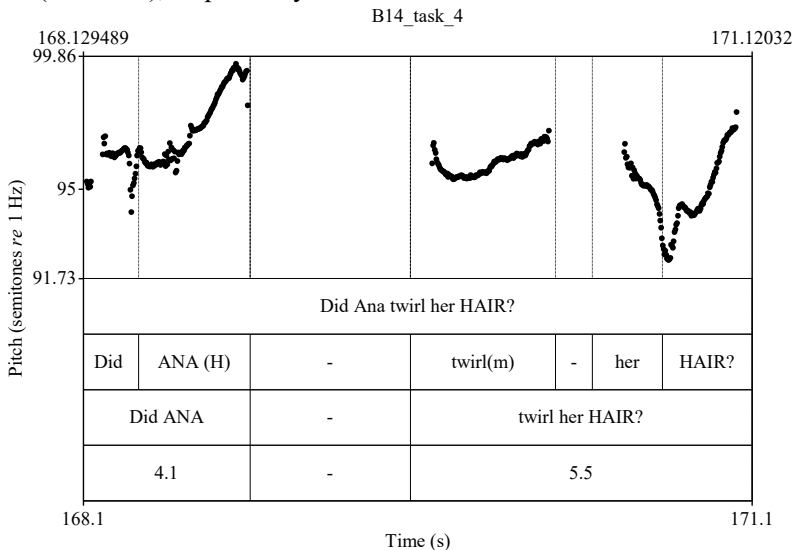


Figure T20: *Did Ana twirl her HAIR?* (Task 4, Item 26) by B14 – An example of an unexpected NS production combined with unexpected syllable lengthening/rising intonation, unexpected pauses, and unexpected pronunciation.

In the utterance in Figure T20, B14 had to nuclear stress *hair* in order to show that she was in doubt if Ana had twirled her hair or her dress. However, B14 produced rising intonation and a pitch change of 4.1 semitones on *Ana* and another rise on *hair* with a pitch change of 5.5 semitones. There was a pause of 717ms between *Ana* and the word *twirl*. The pause after *Ana* and the rising intonation on *Ana* were both probably produced because the speaker was trying to concentrate on how to pronounce the imminent word *twirl*, which sounded like [θru:l], whose pronunciation she was totally unconfident about (rated as 1). The intended NS was placed at the end of the question. To make things yet more confusing for the listener, *Ana* had more duration than *hair* (400 ms). The listener was misled by the pause and thought that B14 meant to emphasize that Ana rather than Maria had twirled her hair, a strange message given the context set for this production.

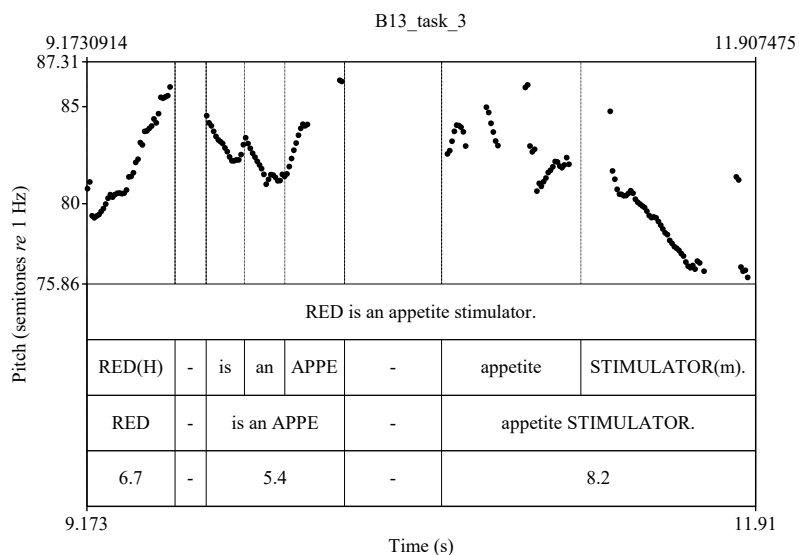


Figure T21: *RED is an appetite stimulator* (Task 3, Item 2) by B13 – An example of an unexpected NS production combined with unexpected rising intonation, unexpected pauses, unexpected pronunciation, and repetition.

In the utterance in Figure T21, B13 had to convey that red rather than blue is an appetite stimulator. Observe that *red* has a rising intonation (H) followed by a short pause of 127 ms. This pause may

have been caused by the presence of challenging words ahead, which made B13 hesitate and correct himself on the pronunciation of *appetite*, relaxing on *stimulator* which was pronounced as ['sɪmjʊleɪtɔr]. Although there was some noise in this production, F13 did not ask for repetition. F13 simply disregarded the pitch changes in the beginning of the sentence (a tendency found in the dataset) and opted for the NS at the final part of the utterance (*stimulator*), which also happened to have (1) the greatest pitch change (8.2 st), probably produced due to the relief of finishing the sentence, and (2) the usual low boundary tone (L%) used in statements.

This appendix discussed the details of unexpected productions combined with pronunciation and other prosody issues (e.g., intonation and pauses). A detailed account of the productions was carried out and examples were provided for the aspects raised:

- (1) expected NS with no pronunciation issues;
- (2) expected NS with pronunciation issues;
- (3) unexpected NS with no pronunciation issues, unexpected NS with no apparent pronunciation issues;
- (4) unexpected NS with apparent pronunciation issues;
- (5) unexpected NS with pauses;
- (6) unexpected NS with apparent pronunciation issues and pauses;
- (7) unexpected NS with pronunciation issues and repetition;
- (8) unexpected NS with pronunciation issues, repetition, and pauses;
- (9) unexpected NS with unexpected syllable lengthening or rising intonation;

- (10) unexpected NS with unexpected syllable lengthening or rising intonation and pauses;
- (11) unexpected NS with syllable lengthening or rising intonation and unexpected pronunciation;
- (12) unexpected productions with syllable lengthening or rising intonation, unexpected pronunciation, and pauses; and
- (13) unexpected productions with syllable lengthening or rising intonation, unexpected pronunciation, pauses, and repetition.

The frequency of these aspects were compared to the presence of challenging words in utterances in Section 4.1.3. and it was found that the combination of a greater number of aspects happened more often in utterances with challenging words, either short or long.



Table U1 puts together the percentage for the frequency that the productions in each aspect (related to pronunciation) were (a) perceived as expected, which led listeners to successfully interpret the speakers' intent, and (b) perceived as unexpected, which led to noise in communication (sites of reduced interpretability) or to real communication breakdowns. Rows labeled "Red. Interp." and "Breakdowns" display the percentages for sites of reduced interpretability and real communication breakdowns, respectively, while the row labeled "T. unexp." displays the sum of sites of reduced interpretability and real communication breakdowns, which together represented unexpected perception.

Interpretability was not affected to a great extent when utterances were (a) produced with unexpected pauses only, with no other pronunciation issues (85.7%) (Aspect 5), followed by those (b) produced as expected (76%) (Aspect 1), (c) produced as expected in terms of NS placement but with pronunciation issues (63.1) (Aspect 2), and finally (d) by those that had unexpected NS placements and unexpected syllable lengthening or rising intonation, unexpected pronunciation and pauses (55.2%) (Aspect 12). Nevertheless, observe that the latter and Aspect 5, both unexpected productions involving pauses, when not perceived as expected led straightly to communication breakdowns, as the listener was not suspicious that the speaker would be wanting to convey a message other than the one the listener got.

The remainder of the aspects showed to be even more harmful to successful communication. Aspects 7 and 10 were the ones that showed to hinder communication to a greater extent (83.3%). They refer to utterances produced as unexpected in terms of NS placement accompanied by other unexpected "processes": the repetition of a word/phrase, or of part of it, and unexpected pronunciation (Aspect 7) and syllable lengthening or rising intonation on words other than the one that should be emphasized, which were many times perceived as being pauses (Aspect 10).

The production of unexpected pauses and unexpected syllable lengthening is a problem in terms of prosody processing, as both pauses and syllable lengthening are cues for thought group limits (Heusinger, 2007) and they may break speech in nonsense units. In contexts in which English is used as a contact language, serious misunderstandings can be caused.

In this dataset, both syllable lengthening and pauses were mostly caused by the imminence of challenging words, which were often pronounced in an unexpected manner. Observe that, in Aspect 7, in which pronunciation issues were present, there were more unexpected perceptions that only represented sights of reduced interpretability, that is, when the utterances were repeated, listeners were able to interpret the speakers' intent as expected (50%). On the other hand, in Aspect 10, when no pronunciation issues were present, real communication breakdowns were more frequent (50%), that is, listeners had no idea that they were misinterpreting the speakers' intent. This result goes in line with studies that found that even speakers' speech with expected delivery of segments may be challenging to understand due to unexpected prosody realization (Galaczi et al., 2017).

This appendix presented the perception results based on the details of unexpected productions combined with pronunciation and other prosody issues (e.g., intonation and pauses). From a comparison drawn about the perception of utterances with the presence of only one aspect or the presence of the combination of two or more aspects done in Section 4.2.4.2, it was found that utterances with more aspects harmed the perception of NS and the speakers' intent interpretation to a greater extent.