Semantic vs. Prosodic Prominence – Pronoun Realisation in spontaneous Mono- and Bilingual English

Marlene Böttcher
Kiel University
mboettcher@isfas.uni-kiel.de

Abstract

The phonological realisation of pronouns (reduced vs. explicit) is conditioned by information structure. The choice of a reduced expression depends on the availability in a language. In stress languages like English PRONs are generally unstressed, while focus allows for stressed forms. Pro-drop languages like Turkish, make a distinction between null and overt pronouns. The availability of two structures in bilinguals is prone to transfer. A recent study on spontaneous intonation reported stressed non-focused pronouns in English (Böttcher & Zerbian, 2020). Their bilingual group comprised speakers from different language background (including pro-drop languages) which is taken up in this study.

The present paper presents an analysis of prosodically prominent PRONs and their relation to contrastive focus in spontaneous English narrations by 4 mono- and 16 bilingual speakers (Russian, Turkish, Greek, German) from the RUEG corpus. Speakers of all groups produced stressed PRONs independent of contrast related to the previously reported aspects of prosodic phrasing. Bilinguals frequently left contrastive pronouns unmarked, avoiding stress realisations, while monolinguals marked contrastive PRONs prosodically. The results are in line with the overall tendency for bilingual speakers’ prosody to follow unmarked structural constraints (eg. phrasing) rather than marked pragmatic constraints (MHD, cf. Zerbian, 2015).

Index Terms: PRONs, prominence, English, prosody, contrast, information structure, bilinguals, markedness differential hypothesis

1. Introduction

The realisation of anaphoric expressions such as pronouns (PRON) is related to information structural categories such as givenness, topic and focus ([1]). While reduced expressions (eg. unstressed/null PRONs) are used in cases of given referents and topic continuation, explicit expressions (eg. full NPs) are used in cases of contrastive focus and topic shift, (eg. [2]). The choice of a reduced expression depends on the availability in a language.

English makes use of unstressed PRONs in the case of topic continuation and allows for stressed PRON realisation in cases of focus and if produced in isolation ([3]). Like English, German and Russian are intonation stress languages with comparable prosodic constraints regarding PRONs. PRONs are frequently unmarked, while prominence on PRONs is related to focus marking, especially contrast (eg. [4], [5]). For German there is evidence for stressed PRONs not in focus from both read and spontaneous speech, indicating that their prosodic realization is determined by additional factors eg. number of syllables and syntactic constituency ([6], [7]).

Turkish and Greek, on the other hand, are pro-drop languages. Both languages use null PRONs in the case of given subject referents. The rich morphology, eg. the verb inflection, then provides the necessary information. Overt PRONs are realized in these languages in the case of contrast and topic shift (eg. [2], [8]).

The availability of two syntactic structures in bilinguals (pro-drop vs. overt PRONs) has been investigated with regard to transfer to the other language. There is evidence from Italian-English, Italian-Spanish and Polish-English bilinguals ([9], [10]) showing increased use of overt PRONs in non-topic shift contexts in the pro-drop language (Italian, Polish). However, the same bilinguals show appropriate use of overt PRONs in their English.

Additionally, research on the prosody of bilingual speakers has shown the prosodic marking of information status to be a vulnerable area for language contact situations ([11]). L2 speakers of different language background fail to deaccent given constituents in their English (eg. [12]). And highly proficient L2 speakers who are capable to produce native like intonation struggle to vary in different speaking styles ([13]).

However, a recent investigation of spontaneous speech [14] has shown to allow for the accentuation of given referents, ie. spontaneous intonation differs from reading intonation. This was confirmed for PRONs in an investigation of spontaneous narrations by both mono- and bilingual speakers of English. [15] found stressed un-focused PRONs in connection with prosodic phrasing. Stressed PRONs were realized in phrasal edge position, ie. if realized at the beginning or the end of an intonation phrase (IP) or as its only constituent. The bilingual group in [15] comprised bilingual speakers of Russian, Turkish and Greek. The different language background of the speakers is taken up in this study.

The bilinguals in [15] as well as in this study are heritage speakers (HS), ie. simultaneous or early successive bilinguals. HS are speakers of a weaker minority or heritage language usually spoken in their families and a dominant majority language of the larger society they are part of (eg. [16]). HS have shown to be sensitive to formality in their heritage language. That is, they reveal little to no stylistic differentiation in eg. the use of honorifics or polite PRONs appropriate to a more formal context ([17]).

The differences regarding the reduced PRONs in the respective language and the previous research on bilinguals and stylistic variation (ie. formality) motivate the following research questions for this study on spontaneous English: In how far differ speakers with different language background
i) in the use i.e. number of overt PRONs in spontaneous speech?

ii) in the prosodic realization of contrastive referents produced as a PRON?

iii) in the realization of prominent PRONs in different situations with different formality?

And if there are differences between speaker groups,

iv) Can these differences be explained by influence from a particular heritage language or generally language contact?

2. Method

This study makes use of a corpus by the research project RUEG ([18]). The data of the corpus comprise four heritage languages (German (D), Russian (R), Turkish (T) and Greek (G)) in two dominant language settings (Germany and the US). To this end spontaneous narrations portraying a car accident were elicited ([19]). Participants were confronted with a video depicting an incident involving three groups of people, as well as several participants within each group which can possibly be contrasted. Participants were then asked to imagine telling the story to a friend (informal situation) and providing a witness report to a police officer (formal situation). Participants also provided written statements in both formalities. For this study only the spoken data was considered. The data is transcribed and searchable for normalization and parts of speech. A prosodic annotation is under construction and only available for part of the data. For the study at hand the version 0.3.0 of the corpus was analyzed.

2.1 The corpus

The RUEG corpus holds narrations by monolingual (mono) and bilingual (bi) speakers with different language backgrounds. From the English language data, a subset of 40 narrations by 20 speakers (4xmono, 4xbiD, 4xbiR, 4xbiT, 4xbiG) was selected. The analyzed corpus consisted of 35 minutes of speech material. The average narration is 53.4 s long and consists of 150 words of which 10 are PRONs. Table 1 shows the distribution of duration, words and PRONs across speaker groups.

Table 1: Averages of seconds (t), words, PRONs, referential (Ref) and deictic (Deix) PRONs in a narration in formal (fs) and informal situation (is) across speaker groups (mono=monolingual, bi=bilingual, R=Russian, T=Turkish, G=Greek)

<table>
<thead>
<tr>
<th>Ø</th>
<th>Duration (s)</th>
<th>Word</th>
<th>Pronoun</th>
<th>Ref</th>
<th>Deix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fs</td>
<td>is</td>
<td>fs</td>
<td>is</td>
<td>fs</td>
</tr>
<tr>
<td>mono</td>
<td>55.0</td>
<td>28.3</td>
<td>148</td>
<td>82</td>
<td>7</td>
</tr>
<tr>
<td>biD</td>
<td>76.0</td>
<td>64.9</td>
<td>189</td>
<td>176</td>
<td>9</td>
</tr>
<tr>
<td>biR</td>
<td>50.8</td>
<td>36.5</td>
<td>114</td>
<td>93</td>
<td>7</td>
</tr>
<tr>
<td>biT</td>
<td>60.0</td>
<td>55.8</td>
<td>176</td>
<td>188</td>
<td>9</td>
</tr>
<tr>
<td>biG</td>
<td>51.8</td>
<td>56.0</td>
<td>155</td>
<td>177</td>
<td>7</td>
</tr>
<tr>
<td>all</td>
<td>58.7</td>
<td>48.1</td>
<td>156</td>
<td>144</td>
<td>8</td>
</tr>
</tbody>
</table>

There is a preference for longer narrations including more words in the formal situation with 58.7 s and 156 words compared to 48.1 s and 144 words in informal situation. The average number of PRONs, however, is higher in informal (N=12) compared to formal situations (N=8). A strong difference between situations in the number of PRONs is produced by Turkish and Greek bilinguals producing more than twice the number of PRONs in informal compared to formal situation (cf. Table 1). The following analysis focuses on the 400 PRONs produced in all 40 narrations.

2.2 Annotation for Information structure

The aim of this research is to look at the prosodic realization of PRONs as expressions of given referents in connection to contrast. Therefore, the narrations were analyzed for information structure. One of the big difficulties when annotating focus and contrast is the affiliation with prosodic prominence (eg.[20]). To resolve this issue, contrast was annotated based on the knowledge of the possible contrasts in the stimuli independently of the audio recording along the four categories in (2, based on the guidelines presented in [21]). The respective referents forming a contrast are underlined while the contrastive referent is given in brackets. In (2a) the dad is selected from and contrasted with the rest of the family. In (2b) and (2c) parallels are expressed between referents, given similar local relations, i.e. across somebody, or similar actions, i.e. dropping something, evoking a contrast on the second referent compared to the first. And finally (2d) illustrates an instance of correction.

(2)

a. selection/partiality: \(\text{(USbi03MT_isE)}\)

\[
\text{there’s this family walking down the sidewalk and } [\text{the dad}]_{\text{contrast}} \text{ was bouncing a soccer ball}
\]

b. semantic parallel: \(\text{(USbi02FR_isE)}\)

\[
\text{um across me were a couple walking uh the woman was pushing a stroller and the man was eh bouncing a ball and then } [\text{them}]_{\text{contrast}} \text{ was a woman um standing}
\]

c. syntactic parallel: \(\text{(USbi03MT_isE)}\)

\[
\text{and as he dropped the soccer ball there’s a female across kind of getting her groceries out the car and } [\text{she}]_{\text{contrast}} \text{ dropped her groceries}
\]

d. replacement/correction: \(\text{(USbi02FT_isE)}\)

\[
\text{there were these two couple uh there was } [\text{this couple}]_{\text{contrast}} \text{ that was coming}
\]

Many PRONs were deictic PRONs (N=163) mostly of the first person. These PRONs are not analyzed for contrast given their similar transcribed surface form in deictic contrastive context and as a discourse marker (eg. I think, you know). The analysis in 3.2 therefore will focus on the remaining 237 referential PRONs of the third person.

2.3 Annotation for Prosody

The 400 PRONs were also annotated for their prosodic realization. The PRONs were considered for their relative prominence which has shown to be a relevant factor in PRON realization in both German ([6],[7]) and English ([15]). Three prominence levels were distinguished: 0, i.e. no prominence, 1, i.e. weak prominence, and 2, i.e. full prominence co-occurring with a pitch accent ([22] adapted for English). While full prominence co-occurs with pitch accents (PA) the weak prominence level captures reduced prominences in the case of e.g. a rhythmically conditioned stress in the prenuclear region. A weak prominence therefore co-occurs with a strong form but not necessarily PA placement. PA types were annotated descriptively given the bilingual inventory of the HS along MAE ToBI principles (eg.[23]). Prosodic phrasing was
analyzed along universal principles following [24]. The prosodic annotation was carried out using an audio-visual analysis using the speech analysis tool Praat ([25]).

3. Results
The majority of the 400 produced PRONs was not prominent (N=270), 91 PRONs were produced with a reduced prominence and 39 PRONs carried full prominence.

3.1 Prominent PRONs across speaker groups and situations
Of the PRONs about one third (N=130, 34%) were produced with prosodic prominence. As can be seen in Table 2, monolinguals produced relatively more prominent PRONs (N=27, 45%) as well as relatively stronger prominent PRONs (N=12, 20%) compared to bilingual speakers.

Table 2: Number and percentages of prosodically prominent PRONs along three prominence levels and PRONs carrying a pitch accent by speaker group (mono=monolingual, bi=bilingual, R=Russian, T=Turkish, G=Greek)

<table>
<thead>
<tr>
<th>Prominence Level</th>
<th>mono</th>
<th>biD</th>
<th>biR</th>
<th>biT</th>
<th>biG</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Prominence</td>
<td>0</td>
<td>45</td>
<td>12</td>
<td>15</td>
<td>42</td>
</tr>
<tr>
<td>Low</td>
<td>1</td>
<td>28</td>
<td>23</td>
<td>19</td>
<td>28</td>
</tr>
<tr>
<td>High</td>
<td>1</td>
<td>14</td>
<td>28</td>
<td>19</td>
<td>22</td>
</tr>
</tbody>
</table>

Regarding the different situations, monolinguals show a tendency for stressed PRONs in formal situations (N=29 (66%); N=8 (23%)). German and Russian bilinguals show a similar tendency, however, to a lower proportion (D: N=14 (40%); N=13 (29%); R: N=9 (33%); N=6 (19%); turquoise and Greek bilinguals produced more prominent PRONs in informal situations (T: N=7; 20%; N=21; 27%; G: N=8; 30%; N=15; 25%, see also Figure 1), yet the number of instances is quite low.

![Figure 1: Accented PRON realizations in percent (absolute numbers are given in the columns) in different situations across speaker groups (mono= monolingual, bi= bilingual, R=Russian, T=Turkish, G=Greek)](image)

3.2 Prosodic marking of contrastive PRONs
Of the 237 referential PRONs produced in all narrations 73 were realized with prosodic prominence while 47 were categorized as one of the contrast categories in (2). There were, however, prominent PRONs not expressing a contrast and prosodically unmarked PRONs expressing a contrast (see Table 3). This gives rise to the question how contrastive PRONs are realized prosodically by the different speaker groups.

Table 3: Number of prosodically and semantically prominent referential PRONs (3rd person)

<table>
<thead>
<tr>
<th>Semantic Prominence</th>
<th>Contrast</th>
<th>No Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>28</td>
<td>45</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
<td>145</td>
</tr>
</tbody>
</table>

3.2.1 Semantically prominent PRONs
Of the 47 contrastive PRONs, 19 were left prosodically unmarked. While monolinguals produced prosodic prominence on such PRONs in 90% of the cases, bilinguals produced several prosodically unmarked contrasts on PRONs (see Table 4). The bilingual Turkish speaker group produced the highest number of prosodically unmarked PRONs expressing a contrast (N=10).

Table 4: Numbers of prosodically prominent PRONs expressing a contrast by speaker group (mono=monolingual, bi=bilingual, R=Russian, T=Turkish, G=Greek)

<table>
<thead>
<tr>
<th>Prominence Level</th>
<th>mono</th>
<th>biD</th>
<th>biR</th>
<th>biT</th>
<th>biG</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Prominence</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Low</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>High</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Pitch accent</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PRONs</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

3.2.2 Prosodic prominence only
Next to these contrastive PRONs there were also 45 referential PRONs not part of a contrastive expression produced with prosodic prominence. They all fall into one of three categories provided in (3) which have been previously reported in [15]. Prominent PRONs are provided in capitals. IPs are indicated by vertical dashes.

(3)

a. Topic and then the cars | (USmo01FE_isE) | [and then the cars] | there were two cars coming up | and | [THEY"D L.~] | [tops] | to stop abruptly ] |

b. Phrase final (USmo03FE_isE) | [the dog] | [had ended up seeing] | the soccer ball | [and ran] | in the middle of the street | [to play with IT H.~ L.~] |

c. Small IP (USbi06MT_isE) and as soon as this guy tries to cross the street | [HE H.~ L.~ ] | [he accidently] | [lets go of the ball] |

In (3a) a topic is produced with a prenuclear accent early on in an IP. The example (3b) shows a phrase final PRON. The spontaneous nature of the data also lead to a number of separately phrased PRONs as in (3c) due to hesitation. These instances of PRONs were produced predominantly by bilinguals and most frequently in phrase final position. (see Table 5).
Table 5: Number of prosodically prominent PRONs expressing no semantic contrast regarding prosodic phrasing by speaker group (mono=monolingual, bi=bilingual, R=Russian, T=Turkish, G=Greek)

<table>
<thead>
<tr>
<th></th>
<th>mono</th>
<th>biD</th>
<th>biR</th>
<th>biT</th>
<th>biG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase final</td>
<td>2</td>
<td>12</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Small IP/hesitation</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Phrase initial/Topic</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

These instances of stressed PRONs were not analyzed as expressions of contrast. However, given the spontaneous nature of the data it is still possible that speakers were having a contrast in mind.

4. Discussion

In the analyzed data the majority of PRONs were indeed prosodically reduced in line with predictions based on the literature (eg. [1], [3]). Yet, PRONs were frequently stressed by speakers of all groups and contrastive PRONs were predominantly marked prosodically. The results for monolinguals suggest that these speakers produce more and also stronger prosodic prominences on PRONs compared to all bilingual speaker groups. Although the referential PRONs in the analysed data are related to given referents, speakers from all groups produce prominent PRONs independent of contrast. These instances are in line with previous findings on the nature of spontaneous speech ([14]). For PRONs the aspects of topic realisations with prenuclear prominence and prosodic phrasing reported in [14] were confirmed in the data analysed here.

Research question i) addressed in how far speakers with different language background differ in the use ie. number of overt PRONs in spontaneous speech. The analyzed data suggests that there is a trend for bilinguals with a pro-drop heritage language to produce more instances of PRONs in the informal situation. For the other speaker groups the number of PRONs across situations is fairly similar.

The speaker groups also differ regarding the realisation of unaccented contrastive PRONs providing insight into research question ii). While English monolingual speakers use prosodic prominence to mark semantic prominence in the form of contrast, bilinguals produce less stress on such contrastive PRONs. They additionally show a difference regarding the two situations answering the third research question iii). While monolinguals accent (almost) all contrastive PRONs independent of the situation, bilinguals realise more accents on such PRONs in informal (69%) compared to formal situations (21%).

An interesting observation regarding research question iv) is the high number of non-prominent but contrastive PRONs by speakers of Turkish. As a pro-drop language, Turkish prefers null PRONs except for contrast in which case the PRON is necessarily realised. One explanation would be that those speakers with a pro-drop heritage language produce PRONs expressing a contrast rather than marking them prosodically. A similar overgeneralization of overt PRON use was observed in the heritage language of such bilingual speakers ([17]). The results of the data analysed here could hint at a tendency for a similar overgeneralization in their dominant language. Yet, both Turkish and Greek are pro-drop languages and the data of Greek bilinguals do not support the hypothesis of a such a pro-drop language contact phenomenon. A closer look at other referent realisations, such as null PRONs, PRONs and full lexical NPs, and their prosodic realisation would be necessary to come to a definite conclusion.

Overall, the small set of data analysed here suggests that bilingual speakers avoid prosodic prominence on PRONs even in cases of contrast while overgeneralizing the de-accentuation of PRONs as given constituents and function words. Their realisation of prominent PRONs rather follows structural constraints of prosodic phrasing. These observations are in line with the Markedness Differential Hypothesis (MDH, [26]), which states that marked structures are more difficult to acquire in language contact than unmarked structures. Adapted for the prosody of English, structural constrains, eg. syntactically conditioned phrasing, are seen as unmarked, while pragmatic constraints, eg. accentuation of contrastively focused constituents, are marked ([27]). The bilinguals in this study produce prosodic prominence following unmarked structures in connection to syntactic and prosodic phrasing. The pragmatic constraints considered here do not always lead to prosodic prominence on PRONs. This is especially surprising given that these bilinguals are adult HS with high proficiency in English as their dominant language. The analysed data, however, suggests that these heritage language speakers show similarities in the prosody of their dominant language to other bilinguals when compared to monolingual speakers. Research comparing HS not only to monolinguals, but also other bilinguals could provide further insight into their emerging grammar. Additionally, investigations of other aspects of HSs’ dominant language are necessary and will be addressed in the further work within our research group RUEG.

5. Acknowledgements

The research results presented in this publication were funded by the German Research Foundation (DFG) as part of the research unit Emerging grammars in language contact situations: a comparative approach (FOR 2537) in project(s) P7 (project no. 394844953 , GZ. ZE 940/2-1). Thank you to the attendees of the research colloquia at department of general linguistics and phonetics at the University of Cologne as well as at the department of English linguistics at the University of Stuttgart for feedback and support.

6. References

Australia 2019 (pp. 2640-44). Canberra, Australia: Australasian Speech Science and Technology Association Inc.


