THE EFFECT OF F0 AND POST-FOCAL COMPRESSION ON THE PERCEPTION OF NARROW FOCUS IN HEBREW

Noam Amir[1], Vered Silber-Varod[2], Ori Anavi[1], Ella Deouell[1], Hansjörg Mixdorff[3]

[1]Department of Communication Disorders, Tel Aviv University
[2]Center for AI and Data Science (TAD), Tel Aviv University
[3]Berliner Hochschule für Technik
noama@tauex.tau.ac.il, other email addresses?

ABSTRACT

In this study, we examine the effect of shifts in fundamental frequency (F0) on the perception of narrow focus (NF). F0 was manipulated systematically on the first word in a subject-verb-object (SVO) sentence, and separately on the rest of the sentence, using the Fujisaki model of intonation. The original sentence was taken from an existing Hebrew corpus where it had previously been found to be uttered with narrow focus and measurable post focal compression (PFC). Findings indicate, as expected, that systematically lowering F0 on the initial word reduced recognition of NF. Systematically raising F0 in the rest of the sentence also reduced recognition of NF. Notably, raising F0 in the second part of the sentence reduced recognition rates, even when no F0 manipulation was performed on the first word. This indicates that PFC is not only a production strategy of Hebrew speakers but has a role in perception of NF also.

Keywords: Narrow Focus, Fujisaki Model, Post Focal Compression.

1. INTRODUCTION

Communication is based not only on words, but also on supra-segmental features that allow us to convey meanings beyond the words alone, such as intentions, feelings, irony and humor. Changes in prominence affect supra-segmental features, and make a part of a sentence (e.g., a word) stand out by changes in pitch, duration and intensity [1, 2, 3] and hence modifies the semantic focus. There are two types of focus, according to the length of the unit of the sentence being highlighted: Broad or “default” Focus (BF), which concerns the whole sentence, and Narrow Focus (NF), which emphasizes, for instance, one word within the sentence [4, 5]. The purpose of the focus is to convey new information to the listener [6]. Thus, the type of focus (broad vs. narrow) and its location in the sentence is based on the information the speaker wants to provide or the question that he or she is answering [7, 5].

In most languages, NF is created by increasing fundamental frequency, duration and intensity of the word being emphasized [8, 9, 10, 11, 12, 13]. There is evidence indicating that in some languages, F0 is the most important feature in the recognition of NF (as reviewed in Terken [14]). There is supporting evidence for this in Hebrew as well [15], though more research needs to be done.

Interestingly, it has been shown that in some languages, including Hebrew [17], acoustic changes do not occur only in the focused word but also in the words following it [16, 11]. Post-focal compression (PFC) is the compression in F0 of the units following the prominent word. Studies show that there are languages in which PFC is crucial to the perception of NF [16]. Evidence of that has been shown in a few studies, as Xu [16] reviewed, which found that NF perception worsens when the word being emphasized is at the end of the sentence, since there are no units after it that can undergo compression. Unpublished studies conducted by the present authors showed the same effect in Hebrew [17]. This indicates that there might be an effect of the PFC on the perception of NF in Hebrew, though this has not been directly examined yet. This was therefore the primary goal of the present study.

Stimuli used to assess the perception of NF can be produced by using computer software, which allows isolating and systematically changing the acoustic features in a chosen word in the sentence [3], such as the FujiParaEditor software based on the Fujisaki model in conjunction with Praat software. The Fujisaki model provides a representation of the F0 contour along the sentence using the superposition of the constant base frequency, the phrase components (manipulated by the phrase commands) and the accent components (manipulated by the accent commands). Neutral sentences are characterized with low and wide accent components while NF sentences are characterized with high and narrow accent components. Studies show that the amplitude of the accent commands manipulated by this model is highly correlated with judge's perception of NF [3].

In the present study, we tried to shed light on the acoustic features contributing to the perception of NF in Hebrew. To this end we employed the Fujisaki model to create systematic changes in F0, separately on the focused word and on the units that follow it.

2. METHODS

A set of 300 three-word target stimuli and 100 non-targeted stimuli were presented to 20 listeners. Their task was to determine whether any one of the three words in each stimulus was emphasized, and to what degree. Further details are presented below.

2.1. Stimuli

The target stimuli were based on a single utterance recorded in a previous production study on narrow focus [17]. The selected utterance achieved a 70% recognition score for focus on the first word. This utterance was a three word sentence in Hebrew: *Xatul nixnas l-a-gan* (cat.SG.M enter.past.SG.M. to-the-garden) "A cat entered the garden". The intonation contour of this utterance was modified using software based on the Fujisaki model of intonation [18]. Analysis of the original intonation contour determined the presence of two accent commands – a strong one on the initial word, and a weaker one on the following two words, as shown in Figure 1.

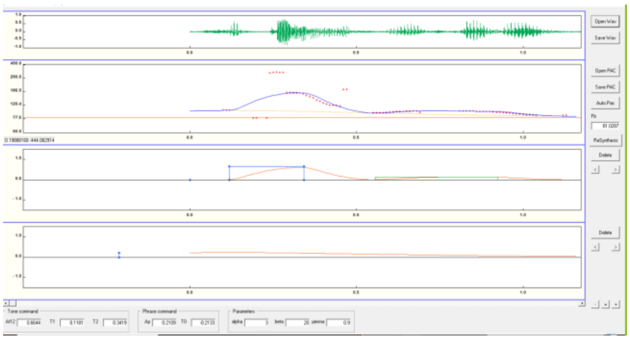

Figure 1: Screenshot from the FujiParaEditor.

The utterance was then resynthesized 30 times as follows: Leaving the second accent command intact, the first accent command was reduced in five increments, in order to lower the pitch peak in the first word in steps of 10 Hz, from its initial value of 176Hz down to 126Hz. Next, the second accent command was raised, to heighten the pitch peak in the second part of the utterance from 100Hz to 140Hz. Once again, six variations were created with six different heights in the first pitch peak. This was repeated five times. Thus, the second pitch peak attained a value of 150Hz in the final iteration. In total this process yielded 30 stimuli, comprising all the different combinations of six different initial peak heights and five different secondary pitch peak heights (Table 1).

Each of the resultant stimuli was presented for judgment 10 times, giving 300 presentations of the target stimuli. To avoid boredom and to prevent the listeners from guessing what manipulations were performed, 100 additional three-word sentences were presented, but not analysed.

Table 1: The different variations of the target stimulus.

F0 peak associated with the first accent component (F0_NF)	F0 peak associated with the second accent component (F0_PFC)
F0_NF1=176Hz	F0_PFC1=100Hz F0_PFC2=110Hz F0_PFC3=120Hz F0_PFC4=130Hz F0_PFC5=140Hz
F0_NF2=166Hz	F0_PFC1=100Hz F0_PFC2=110Hz F0_PFC3=120Hz F0_PFC4=130Hz F0_PFC5=140Hz
F0_NF3=156Hz	F0_PFC1=100Hz F0_PFC2=110Hz F0_PFC3=120Hz F0_PFC4=130Hz F0_PFC5=140Hz
F0_NF4=146Hz	F0_PFC1=100Hz F0_PFC2=110Hz F0_PFC3=120Hz F0_PFC4=130Hz F0_PFC5=140Hz
F0_NF5=136Hz	F0_PFC1=100Hz F0_PFC2=110Hz F0_PFC3=120Hz F0_PFC4=130Hz F0_PFC5=140Hz
F0_NF6=126Hz	F0_PFC1=100Hz F0_PFC2=110Hz F0_PFC3=120Hz F0_PFC4=130Hz F0_PFC5=140Hz

2.2. Listeners

Twenty female listeners participated in the study (M=27.4, STD=2.5), aged 20 to 31.

2.3. Procedure

Stimuli were presented to the listeners using custom written software employing a graphic user interface (GUI). After hearing each stimulus, the listeners could mark one of four radio buttons, indicating that word 1, 2 or 3 were emphasized, or that no word was emphasized. If options 1-3 were selected, listeners were asked to rate the degree of emphasis on a 5-point scale from weakly emphasized to highly emphasized (The latter was not further analysed in the present study). Stimuli were presented in a random order. Since each distinct stimulus was presented to each subject 10 times, each stimulus received a recognition score of 0 to 100%, from each individual listener.

3. RESULTS

Plotting overall mean percent-correct scores vs both manipulations, gave a 3D plot shown in Figure 2.

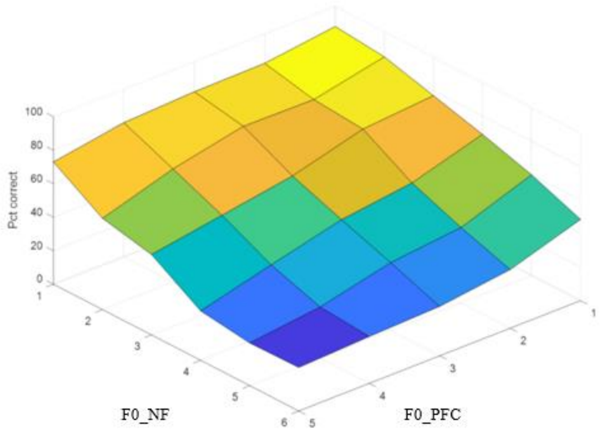

Figure 2: Mean recognition percentages of narrow focus as a function of both manipulations.

The plot clearly suggests that both manipulations affected recognition of narrow focus. It demonstrates that manipulation of the focused word alone has a strong effect, reducing mean recognition from nearly 100% in the unmanipulated condition to about 40% when manipulation was strongest (condition F0_NF6). Manipulation of PFC had a smaller effect, reducing mean recognition to about 70% in the most extreme manipulation (condition F0_PFC5). The plot also suggests that the two manipulations interact, so that combining a manipulation of the focused word with manipulation of PFC reduced recognition more than each manipulation alone. An exploratory ANOVA with repeated measures was therefore conducted with two within-subject factors: narrow focus manipulation (F0_NF, with 1 indicating no manipulation, up to 6 indicating a **reduction** of 50 Hz in main peak height) and post-focal-compression manipulation (F0_PFC) with 1 indicating no manipulation, up to 5 indicating an **increase** of 40 Hz in secondary peak height). Significant main effects were found for both F0_NF ($F(1.58,28.39)=47.11$, $p<0.001$) and F0_PFC ($F(1.46, 26.33)=28.64$, $p<0.001$) and their interaction ($F(20,360)=2.25$, $p=0.002$). This indicates that both factors affected narrow focus recognition significantly. The significant interaction, however, indicates that F0_PFC did not affect recognition in the same manner for different levels of F0_NF. This can be seen in the interaction plot in Figure 3, showing separate lines for the different values of F0_NF, with the x-axis indicating different levels of F0_PFC.

We wished to specifically examine whether PFC did indeed influence recognition scores at all levels of F0_NF, therefore we further conducted six ANOVAs with F0_PFC as a within subject factor, for each level of F0_NF separately. In effect, this is equivalent to examining each line in Figure 3 separately. A highly significant main effect was found in all six, as well as significant linear contrasts, as summarized in tables 2 and 3. This indicates that PFC has a significant effect on recognition of NF, with recognition rates decreasing linearly as PFC diminishes (i.e. as F0 increases on words following the focused word).

Table 2: significances of main effects for the six ANOVAs.

	df	*F*	*p*
F0_NF1	[4,76]	6.46	0.000
F0_NF2	[4,76]	14.43	0.000
F0_NF3	[4,76]	10.34	0.000
F0_NF4	[4,76]	14.12	0.000
F0_NF5	[4,76]	14.40	0.000
F0_NF6	[4,76]	7.04	0.001

Table 3: significances of linear contrasts for the six ANOVAs.

	df	*F*	*p*
F0_NF1	[1,19]	23.65	0.000
F0_NF2	[1,19]	31.38	0.000
F0_NF3	[1,19]	21.88	0.000
F0_NF4	[1,19]	20.40	0.000
F0_NF5	[1,19]	26.48	0.000
F0_NF6	[1,19]	10.76	0.004

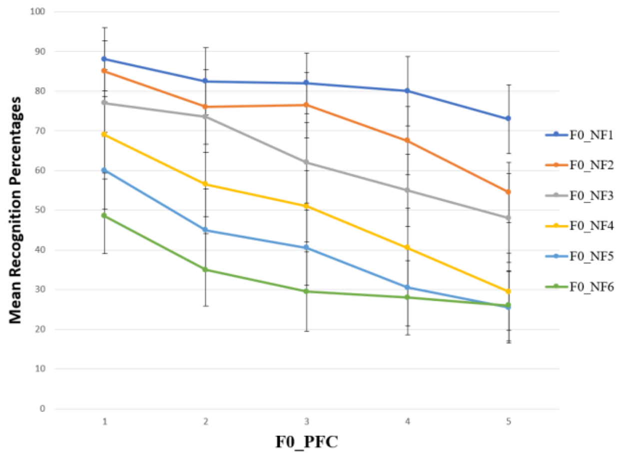

Figure 3: Mean recognition percentages of the narrow focus at the different levels of F0_PFC. Each line represents a different value of F0_NF.

4. DISCUSSION AND CONCLUSIONS

The results of this study highlight the role of F0 in creating the perception of narrow focus in Hebrew. As expected, the height of the pitch peak on the focused word has a major influence on narrow focus perception. PFC, however, influences this perception also. Even when the peak on the focused word is at its highest, reducing PFC (i.e. increasing the secondary pitch peak) reduces perception scores of narrow focus. This indicates that PFC is not only a production strategy, but that listeners rely on this cue also in judging NF.

Furthermore, the results indicate that narrow focus is not a categorical phenomenon in Hebrew. Gradual decrease in pitch on the focused word, and gradual decrease of PFC, resulted in gradual changes of NF recognition scores, in contrast to reports in the literature on other languages [7].

A future publication on data we have already collected [17] will address production strategies, demonstrating that PFC is indeed widely employed by speakers of modern Hebrew.

6. REFERENCES

[1] Raphael, L. J., Borden, G. J., Harris, K. S. 2007. *Speech science primer: Physiology, acoustics, and perception of speech,* 5th ed. Lippincott Williams & Wilkins.
[2] Lehiste I. 1970. *Suprasegmentals*. Cambridge, MIT, Press.
[3] Rao, P., Sanghvi, N., Mixdorff, H., Sabu, K. 2017. Acoustic correlates of focus in Marathi: Production and perception. *Journal of Phonetics*, 65, 110–125.
[4] Vilaplana, E. E., Maidment, J. 1999. Central Catalan declaratives: the relation between focus and downstep. *14th International Congress of Phonetic Sciences*, San Francisco, CA, United States.
[5] Sityaev, D., House, J. 2003. Phonetic and phonological correlates of broad, narrow and contrastive focus in English. *The 15th International Congress of Phonetic Sciences*, Barcelona, Spain.
[6] Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R., Lotts, D. W. 1986. Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments. *Language and Speech*, 29(3), 233–251.
[7] Ladd Jr, D. R. 1980. *The structure of intonational meaning: Evidence from English*. Indiana University Press.
[8] Breen, M., Fedorenko, E., Wagner, M., Gibson, E. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes*, 25(7–9), 1044–1098.
[9] Hanssen, J. E. G., Peters, J., Gussenhoven, C. 2008. Prosodic effects of focus in Dutch declaratives. *Speech Prosody 4thInternational Conference*, Campinas, Brazil.
[10] Lee, A. Xu, Y. 2012. Revisiting focus prosody in Japanese. *Speech Prosody 6th International Conference*, Shanghi, China.
[11] Mixdorff, H., Amir, N. 2002. The Prosody of Modern Hebrew-A Quantitative Study. *Speech Prosody International Conference*, Aix-en-Provence, France.
[12] Mixdorff, H., Cossio-Mercado, C., Hönemann, A., Gurlekian, J., Evin, D., Torres, H. 2015. Acoustic correlates of perceived syllable prominence in German. *16th Annual Conference of the International Speech Communication Association* (INTERSPEECH 2015), Dresden, Germany.
[13] Xu, Y., Xu, C. X. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, 33(2), 159-197.
[14] Terken, J. 1991. Fundamental frequency and perceived prominence of accented syllables. *The Journal of the Acoustical Society of America*, 89(4), 1768–1776.
[15] Amir, N., Almogi, B. C., Mixdorff, H. 2008. A systematic framework for studying the contribution of F0 and duration to the perception of accented words in Hebrew. *Speech Prosody 4th International Conference*, Campinas, Brazil.
[16] Xu, Y. 2011. Post-focus Compression: Cross-linguistic Distribution and Historical Origin [Paper presentation]. *17th International Congress of Phonetic Sciences* (ICPhS 2011), Hong Kong, China.
[17] Bechar Marom, N., Hechter, N., Tahayu, Y., Amir, N., Silber-Varod, V. 2021. Prosodic focus in Hebrew: A Perception study. Proceedings of Tone and Intonation 2021 (TAI 2021), *The 1st Internaional conference on TAI* (pp. 192–193), Sønderborg, Denmark.
[18] Fujisaki, H. and Hirose, K. 1984. Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *Journal of the Acoustical Society of Japan*, 5(4): pp. 233–241.