

Non-Native Perception of Native English Prominence

Rachel E. Baker¹

¹Department of Linguistics, Northwestern University, Evanston, IL, USA

r-baker2@northwestern.edu

Abstract

This study examines the factors that influence Mandarin and Korean speakers' understanding of English pitch accent placement. The results show that both groups have difficulty determining whether a native English speaker is using the correct prosody for a particular discourse context. Participants' ability to do this task improved with greater English proficiency, as measured by the Versant English Test. Participants' success was also influenced by whether a sentence had broad or narrow focus, whether an accent was placed within a focused constituent, and by the number of contexts in which a particular accent placement is used.

Index Terms: cross-language, pitch accent, perception

1. Introduction and Background

1.1. Introduction

Prosodically prominent words in English provide clues to the location of focus in an utterance. Focused words and phrases are new or otherwise informative. Although there is a relationship between focus and prominence in English, not all words in a focused phrase must be prominent, and some prominent words are not in focus [1]. This may make it hard for English learners to interpret patterns of prominence.

Adding to this difficulty is the diversity of focus-marking strategies that exist cross-linguistically. Languages use prosodic, syntactic, and morphological methods to mark focus [2]. Even languages that use prosody to mark focus can differ in how they mark focus on larger constituents, such as verb phrases and sentences. In these cases, some part of the focused constituent may be prosodically prominent, or focus may not be marked at all. Different languages can also make different words prominent within these larger focused constituents.

It is important for English-learners to accurately interpret English prosodic prominence, so they can fully understand English utterances and improve their production of English prosody. Despite its importance, little work has examined non-native understanding of English phrase-level prosody. This study examines native Mandarin and Korean speakers' ability to determine whether a native English speaker is using the correct prosodic prominence for a particular discourse context.

1.2. English Focus Marking

The autosegmental-metrical prosodic framework proposes that English speakers use pitch accents to make words prominent [3]. Pitch accents are local pitch targets associated with particular syllables [4]. In English, pitch accents are associated with stressed syllables, which can have increased duration and intensity [5]. English pitch accents are assigned post-lexically, based on the location of focus in a sentence and English-specific accent assignment rules [1].

While a phrase can contain several pitch accents, only the nuclear pitch accent, which is the final pitch accent in an

English prosodic phrase, is obligatory. This nuclear pitch accent is the only pitch accent associated with focus [6]. Prenuclear accents are optional, and their use is determined by the metrical structure of the sentence, rather than its information structure [1]. There are several different types of English pitch accents, each consisting of a high tone (H), a low tone (L), or a combination of the two. English speakers use different pitch accent types to express different meanings [7].

The location of focus in an English sentence partially determines nuclear pitch accent location [1]. In the simplest case ('narrow focus'), only one word is in focus, and it receives a nuclear pitch accent (1).

(1) Q: Who flew home?

A: [JOHN] flew home.¹

In (1A), *John* is in focus because it is not stated in or derivable from the preceding context, it could be said to contrast with all other possible answers [8], and it corresponds to the wh-word in the preceding question.

When an entire phrase is in focus ('broad focus'), only one word within it receives a nuclear pitch accent (2).

(2) Q: What did John do?

A: John [flew HOME.]

In (2A), the verb phrase (VP) *flew home* is in focus. However, only the object (*home*) is accented. This is due to a tendency in English to accent arguments over heads (captured by Schwarzschild's [9] constraint stating that heads are less prominent than their internal arguments).

A similar situation can arise when an entire sentence provides new information, and is therefore in broad focus (3).

(3) Q: What happened?

A: [John flew HOME.]

Once again, only the object receives a nuclear pitch accent in (3A).

1.3. Mandarin and Korean Focus Marking

Mandarin is a tone language, which means that pitch targets are assigned to words in the lexicon. Focus can be marked in Mandarin both syntactically and prosodically. Words that are in informational focus can be placed in sentence-final position, and words that are in contrastive focus can be surrounded by focus markers [10]. Mandarin, unlike English, does not require that phrases have any prosodically prominent word, and this is generally the case when a sentence has broad focus or when narrow focus is marked syntactically [10]. Narrow contrastive focus in Mandarin can be prosodically marked by an increased duration and F0 range for the focused word, and a lower maximum F0 and reduced F0 range for all following words [11][12]. VP broad focus sentences are produced with no extra prominence on any word in the VP [13].

¹ In these examples, capitalization indicates the location of the nuclear pitch accent. Square brackets indicate the location of focus, given the context.

Focus in Korean is marked by a word's position in a phrase, with new information appearing at the beginning of a phrase [14]. Korean words in narrow contrastive focus have longer first syllables, higher peak F0s, and greater F0 ranges [15]. In the post-focal area, words are shorter, and Accentual Phrase (AP) boundaries are often removed, leading to a loss of boundary tones associated with the edges of the APs [15]. VP broad focus is also prosodically marked in Korean. An Intonation Phrase boundary is inserted before the focused VP, leading to a boundary tone at the end of the subject (preceding the VP) [16]. Acoustically, the first word in the focused VP has an expanded pitch range and duration [16].

1.4. Prosody Perception in Non-Native English

The field of non-native prosody perception is relatively new. The work that has been done shows that listeners have difficulty processing prosody in their non-native language. For instance, Akker and Cutler [17] found differences in the speed with which native English speakers and Dutch learners of English process accented and focused words. However, this study did not examine non-native speakers' ability to understand the relationship between accent location and focus.

The goal of the current study is to identify factors that make it hard for language learners to interpret prosodic focus marking in their non-native language. Such factors can relate to differences between a language learner's native language and the target language, or to features of the target language itself. If a language learner's native language affects their interpretation of target language prosody, we would expect to see differences between people with different native languages. For example, Mandarin speakers may have more difficulty than Korean speakers in interpreting prosodic marking of broad focus in English, because they do not mark broad focus prosodically in their native language. Aspects of the target language's focus-marking system may also play a role in non-native prosody interpretation. One factor relating to the target language itself is the relationship between focus and prominence. For instance, English narrow focus may be easier to understand than broad focus, because in narrow focus there is a direct correspondence between focus and accent placement. Another possible target language factor is the frequency with which certain accent patterns are used in the target language. If a certain pattern is used in multiple contexts, non-native speakers may apply it too broadly. These hypotheses are tested using a perception task in which English learners decide if a sentence's prosody is appropriate for its context.

2. Methods

2.1. Participants

Forty non-native English speakers participated in this experiment. Twenty were native Korean speakers (14 female, 6 male), and twenty were native Mandarin speakers (13 female, 7 male). Twenty native English-speaking controls also participated (14 female, 6 male). All of the non-native English speakers were living in the U.S. at the time of the experiment. They had all first moved to an English-speaking country when they were 17 or older, and had lived in English-speaking countries for less than six years.

After completing the experiment, the non-native participants took the Versant English Test (www.ordinate.com/products/english.jsp), an oral test of general English proficiency. The test assesses non-native

English speakers' sentence mastery, vocabulary, fluency, and pronunciation by having them read aloud, repeat sentences, answer questions, rearrange phrases to form sentences, and retell a story. There was no significant difference between the Versant scores for the Korean and Mandarin groups ($U=230$, $p=0.42$).

2.2. Materials

Twenty-four subject-verb-object sentences were constructed. A female native English speaker was recorded reading the sentences in three contexts. In the subject narrow focus (SuNF) contexts, the sentence was preceded by a question about its subject (e.g. Q: Who bought a fan?, A: Kim bought a fan.). In the VP broad focus (VPBF) contexts, the sentence was preceded by a question about its VP (e.g. Q: What did Kim do?, A: Kim bought a fan.). In the sentence broad focus (SBF) contexts, the sentence was preceded by the question "What happened?". A second female native English speaker was recorded reading the context questions.

2.3. Experiment Design

In the instructions, participants were trained on the meaning of the word 'prosody' using descriptions and recorded examples. The example context in the instructions put narrow focus on the verb (a context type that was not used in the experiment). During the experiment, participants saw question-answer pairs written on the screen in standard orthography, with no indication of prominence. They clicked on a button to hear a recording of the question and the answer. They then had to answer the question "Is the prosody of the answer appropriate given the question?". Participants heard 72 items in all (each of the 24 sentences in each of the three contexts). The items were presented in pseudo-random order.

For half of the items, the question and answer had been recorded together (matched), and for the other half, the answer had been recorded with a different question, so the wrong word was accented in the answer (mismatched). The SBF and VPBF mismatched items used answers produced in SuNF contexts. These are mismatched because a sentence produced in an SBF or VPBF context has a nuclear pitch accent on the object, but a sentence produced in an SuNF context has a nuclear pitch accent on the subject. Similarly, the SuNF mismatched items used answers produced in SBF contexts.

3. Results

3.1 Analysis

Participants' responses were analyzed using linear mixed-effects logistic regression models. This type of analysis avoids spurious effects that can arise when proportion data are analysed using traditional ANOVAs [18]. All regressions included participants and items as random variables. The dependent variable in all regressions was item accuracy (correct or incorrect). For categorical variables, one or more levels were compared to a baseline level. The results include estimates of the coefficients associated with each independent variable, in log odds. For categorical variables, a positive estimate means that participants were more likely to respond correctly to target items than baseline items. For continuous variables, a positive estimate means that higher values of the variable were associated with more correct answers.

Two types of regressions were run on the data: native/non-native and non-native. The native/non-native regressions

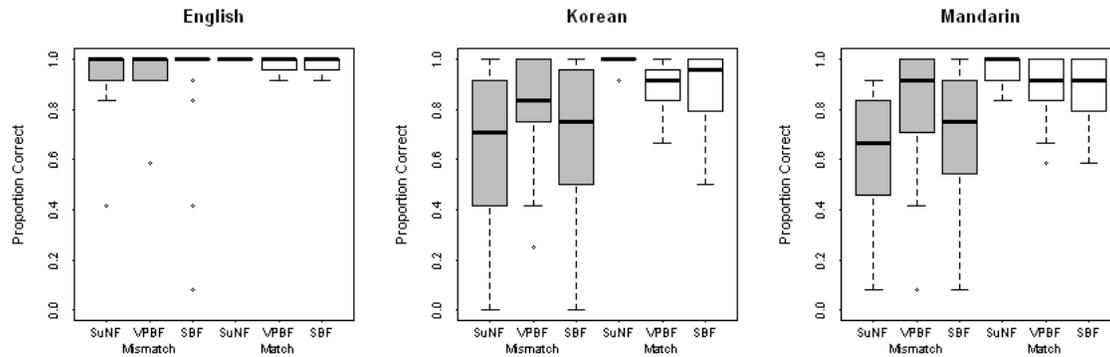


Figure 1. Proportions of correct answer for mismatched and matched items in the three context conditions for the three language groups.

included all three language groups, and compared native English speakers to Mandarin and Korean speakers. The non-native regressions included only the Mandarin and Korean groups. They compared Mandarin speakers to Korean speakers, and controlled for proficiency by adding a Versant score variable. Figure 1 shows the proportions of correct answers for each item type in each language group.

3.2. Native/Non-Native Regressions

The independent variables were language group (English, Mandarin, Korean), match (matched, mismatched), and context (SuNF, VPBF, SBF). For the language group variable, English served as the baseline, and for the match variable, the matched items served as the baseline. Two native/non-native regressions were run, one with the SBF context as the baseline, and the other with the VPBF context as the baseline.

Both native/non-native regressions show that the English group was significantly more accurate than the Korean and Mandarin groups. In addition, the regression with SBF as the baseline showed that accuracy on matched items was significantly higher than accuracy on mismatched items. The significant results for these regressions are listed in Table 1.

	Estimate	Std. Error	z-value	p
SBF Baseline				
Mismatched	-1.637	0.535	-3.063	<0.005
Korean	-2.123	0.624	-3.404	<0.001
Mandarin	-2.189	0.621	-3.525	<0.001
VPBF Baseline				
Korean	-2.076	0.625	-3.323	<0.001
Mandarin	-2.350	0.618	-3.804	<0.001

Table 1. Parameter values for significant effects in the native/non-native regressions

3.3. Non-Native Regressions

The variables in the non-native regressions were the same as those in the native/non-native regressions, except that participants' Versant scores were added as an independent variable, and the variable language group included only the Mandarin and Korean groups, with the Korean group serving as the baseline.

These regressions show that non-natives were significantly more accurate on matched items than mismatched items. Versant score was also a significant predictor of accuracy;

participants with higher Versant scores tended to be more accurate. Non-natives were significantly more accurate on SuNF items than SBF and VPBF items. Finally, there were interactions between match and all pairs of context types. The significant results for these regressions are listed in Table 2.

	Estimate	Std. Error	z-value	p
SBF Baseline				
Mismatched	-1.463	0.269	-5.432	<.001
Versant	0.035	0.011	3.214	<0.005
SuNF	2.400	0.633	3.790	<0.001
Mismatched:SuNF	-2.658	0.672	-3.954	<0.001
Mismatched:VPBF	0.881	0.399	2.205	<0.05
VPBF Baseline				
Mismatched	-0.582	0.291	-2.002	<0.05
Versant	0.035	0.011	3.214	<0.005
SuNF	2.335	0.634	3.682	<0.001
Mismatched:SBF	-0.881	0.399	-2.205	<0.05
Mismatched:SuNF	-3.539	0.682	-5.186	<0.001

Table 2. Parameter values for significant effects in the non-native regressions

To further explore the interactions between match and context, non-native regressions were run on the matched and mismatched items separately. These regressions show that Versant score was a significant predictor of accuracy for both matched and mismatched items. For the matched items, non-natives were significantly more accurate on SuNF items than SBF and VPBF items. For the mismatched items, non-natives were significantly more accurate on VPBF items than SuNF and SBF items. The significant results for the regressions on matched items are listed in Table 3, and those for the regressions on mismatched items are listed in Table 4.

	Estimate	Std. Error	z-value	p
SBF Baseline				
Versant	0.026	0.011	2.316	<0.05
SuNF	2.326	0.650	3.581	<0.001
VPBF Baseline				
Versant	0.026	0.011	2.316	<0.05
SuNF	2.349	0.651	3.609	<0.001

Table 3. Parameter values for significant effects in the non-native regressions on matched items

	Estimate	Std. Error	z-value	p
SBF Baseline				
Versant	0.045	0.016	2.847	<0.005
VPBF	1.016	0.267	3.804	<0.001
VPBF Baseline				
Versant	0.045	0.016	2.847	<0.005
SBF	-1.017	0.267	-3.804	<0.001
SuNF	-1.340	0.265	-5.051	<0.001

Table 4. Parameter values for significant effects in the non-native regressions on mismatched items

4. Discussion and Conclusions

These results show that non-native English speakers struggle to match pitch accent location with intended focus, when listening to native English speech. A non-native English speaker's ability to correctly interpret English pitch accent placement varies with their proficiency. It is also influenced by features of the English focus-marking system, particularly by the number of contexts in which an accent placement is used, whether an accent was placed within a focused constituent, and whether a sentence had broad or narrow focus.

When asked whether the prosody of a sentence was appropriate, non-native participants were overly lenient. They were more likely to correctly accept prosody as appropriate for matched items, than to correctly reject it for mismatched items. This contrasts with the native participants, who had similar accuracy for matched and mismatched items. However, non-natives' performance on this task did improve with greater English proficiency. Participants with higher Versant scores were more accurate than those with lower scores.

Non-native participants had particular difficulty identifying the prosody of a sentence as inappropriate in the SuNF and SBF contexts with mismatched sentences. In the SuNF context, the question asks about the subject of the sentence, but for the mismatched items, the nuclear pitch accent is on the object instead of the subject (4).

(4) Q: Who bought a fan?

A: [Kim] bought a FAN.

Non-natives may incorrectly accept this accent placement because having a nuclear pitch accent on the object is common in English (it is used in SBF, VPBF, and object narrow focus contexts). They could be over-extending this pattern to contexts where it is not appropriate (e.g. the SuNF context). In the SBF context, the question "What happened?" puts the whole sentence in focus, but for the mismatched items, the nuclear pitch accent is on the subject instead of the object (5).

(5) Q: What happened?

A: [KIM] bought a fan].

In this case, the word with the nuclear pitch accent is within the focused constituent (the whole sentence). Non-natives may more easily accept nuclear pitch accents when they are within the focused constituent, even if they are not placed in the standard location. The non-natives were more successful at rejecting inappropriate prosody in the VPBF context. In this context, the question asks about what somebody did, but for the mismatched items, the nuclear pitch accent is on the subject instead of the object (6).

(6) Q: What did Kim do?

A: KIM [bought a fan].

For these items, the nuclear pitch accent is not in a common location or within the focused constituent, so they may have seemed more clearly incorrect.

For the matched items, non-native participants were

especially good at identifying appropriate prosody in the SuNF context. In these items, the question asks about the subject, and the nuclear pitch accent is on the subject (7).

(7) Q: Who bought a fan?

A: [KIM] bought a fan.

Narrow focus items like (7) have the most straightforward relationship between focus and accent location: Only one word is in focus, and it receives the nuclear pitch accent. This makes such items easier to interpret.

These results do not provide evidence for transfer of focus-marking features from a participant's native language to their non-native language, as there were no significant differences between the Korean and Mandarin language groups. However, the non-native participants were most accurate at identifying appropriate prosody in the SuNF context, which is the context that is most similar across the three languages. It may be that Korean and Mandarin are too similar to cause differences in behavior in this type of task. The next step is to examine this issue with a wider variety of native and non-native languages.

5. Acknowledgements

I thank Ann Bradlow, Janet Pierrehumbert, and Brady Clark for their guidance, and Chun Chan for his technical assistance.

6. References

- [1] Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge, UK: Cambridge University Press.
- [2] Büring, Daniel. 2006. Towards a typology of focus realization. Ms., UCLA.
- [3] Pierrehumbert, Janet. 1980. The phonology and phonetics of English intonation. Cambridge, MA: MIT dissertation.
- [4] Beckman, Mary & Janet Pierrehumbert. 1986. Intonational structure in English and Japanese. *Phonology Yearbook* 3. 255-310.
- [5] Beckman, Mary. 1986. *Stress and non-stress accent*. Dordrecht: Foris.
- [6] Welby, Pauline. 2003. Effects of pitch accent type and status on focus projection. *Language and Speech* 46(1). 53-81.
- [7] Büring, Daniel. 2003. On D-Trees, beans, and B-Accents. *Linguistics and Philosophy* 26(5). 511-545.
- [8] Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75-116.
- [9] Schwarzschild, Roger. 1999. GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7(2). 141-177.
- [10] Xu, Liejiong. 2004. Manifestation of informational focus. *Lingua* 114(3). 277-299.
- [11] Xu, Yi, Ching X. Xu & Xuejing Sun. 2004. On the temporal domain of focus. *Speech Prosody*. 81-84.
- [12] Wang, Bei & Yi Xu. 2006. Prosodic encoding of topic and focus in Mandarin. *Speech Prosody*. PS3-12_0172
- [13] Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27. 55-105.
- [14] Jun, Sun-Ah. 2005. Prosody in sentence processing: Korean vs. English. *UCLA Working Papers in Phonetics*. 104. 26-45.
- [15] Jun, Sun-Ah & Hyuck-Joon Lee. 1998. Phonetic and phonological markers of contrastive focus in Korean. *Fifth International Conference on Spoken Language Processing*. 1295-1298.
- [16] Jun, Sun-Ah, Hee-Sun Kim, Hyuck-Joon Lee & Jong-Bok Kim. 2006. An experimental study on the effect of argument structure on VP focus. *Korean Linguistics* 13. 89-112.
- [17] Akker, Evelien & Anne Cutler. 2003. Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition* 6(2). 81-96.
- [18] Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59. 434-446.