It’s Not So Bad: An English Tone for ‘Downplaying’

John C. Wakefield

Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University

j.wakefield@polyu.edu.hk

Abstract

Cantonese has a complex lexical tone system that severely restricts its use of intonation. To compensate for this restriction, sentence-final particles (SFPs) are used to express meanings that are expressed through intonation in other languages. This study examined the pitch contours of Cantonese-to-English translations of the SFP zel in order to test whether the meaning of zel is expressed as a tone in English, and if so, what form this tone takes. Several authors have said that zel expresses the meaning “only” plus a downplaying of the situation.

Based on the F0s of the translations it is concluded that English has a near semantic equivalent of zel, which is expressed by the word “only” plus a rise-fall-rise pitch contour. Guided by the meaning of zel, a definition is proposed for its intonational counterpart in English. Based on previous arguments from the literature, it is proposed that this pitch contour is a floating tone that exists as a morpheme in the English lexicon.

Index Terms: English, Cantonese discourse particles, discourse intonation, floating tone

1. Introduction

Cantonese has a complex lexical tone system that severely restricts its use of intonation. As a result, Cantonese has developed an inventory of more than thirty sentence-final particles (SFPs) to express the types of meanings that are expressed through intonation in other languages (e.g., [2], [4], [13], [17], [18]). SFPs are bound morphemes that attach to the ends of sentences. Throughout the literature on SFPs, linguists have repeatedly compared their functions to those of intonation. It is therefore surprising that only a few linguists have made specific comparisons between given SFPs and given forms of intonation.

One such comparison came from Baker and Ho [1], who said that the way to express the meaning of the question-forming SFP mel in English is to “raise your voice almost to a squeak” (p. 40). Referring to the same particle, Chan [3] said mel could be expressed in English intonationally as “an echo question” (p. 59). Providing empirical evidence for such claims, Wakefield [15] demonstrated that four Cantonese SFPs (lo1, aalmaa3, mel, and aa4) translate quite consistently into English as specific forms of intonation, regardless of the discourse context.

In an attempt to search for yet another English tonal equivalent of an SFP, this study examined the pitch contours of Cantonese-to-English translations of the SFP zel, which is a focus particle that means “only” plus an additional connotative meaning that downplays the situation ([6], [14]). Based on the F0 of the translations, it is concluded that English has a semantic equivalent of zel which is expressed in the form of the word “only” plus a rise-fall-rise pitch contour. Based on arguments from Hirst [9] and Wakefield [15], it is proposed that this pitch contour is a floating tone that exists as a morpheme in the English lexicon.

2. Cantonese “only” SFPs

Fung [6] (pp. 30-73) discussed seven SFPs that all have the unaspirated affricate onset /z/. The literature says that a common feature of all of these particles is the meaning “only.” Fung proposed that each of these particles stems from the core semantic feature “restriction.” This study relates only to the SFP zel, but it is worth comparing it to the semantically simpler SFP zaa3. Fung [6] said that both zaa3 and zel have the semantic features of “restriction” and “exclusion,” but that zel additionally has the semantic feature of “downplaying.” Other authors have said basically the same (e.g., [13], [14]). Kwok [11] said that zaa3 indicates something is insufficient, while zel indicates it is not excessive. This is demonstrated in the following contrast (modified from Fung [6], pp. 59-60):

(1) [In answer to the question “How much is it?”]
   a. Sap man zaa3.
   “It’s only ten dollars.” (less than I thought)
   b. Sap man zel.
   “It’s only ten dollars.” (not enough to matter)

The zaa3-suffixed response in 1a. has the neutral meaning of “only,” and the implication that this is less than the speaker expected is pragmatically understood. This does not support Kwok’s [11] claim that zaa3 implies the amount is “insufficient,” which is a meaning that would be implied if 1a. was said in response to the question “How much money do you have?” In that case the response in 1a. would translate as “(I only have) ten dollars,” implying that it is not as much as the speaker wishes s/he had. Sybesma and Li [14] said everyone seems to “agree that zaa3 conveys ‘only’ in the neutral sense of ‘not more than that’ or ‘and not something else as well’” (p. 1754). It appears that the implication of an “insufficient” amount is context dependent. The reason for using a morpheme with the neutral meaning of “only” (i.e., whether to indicate that an amount is less than expected, insufficient, or something else) will be understood pragmatically from the context in which it is said.

In contrast to 1a. the use of zel in 1b. downplays the amount, meaning that it is not excessive, i.e., that it is not an amount sufficient enough to be considered as a waste of one’s money. This appears to be an intrinsic meaning of zel rather
than something that is context dependent. \textit{Ze1} is used when the speaker believes that the listener holds a particular stance regarding the severity of a particular situation. The contrast between \textit{zaa3} and \textit{ze1} is expressed in the following proposed definitions:

\begin{align}
(2) \quad \text{\textit{zaa3}} &= \text{it is not more than this} \\
(3) \quad \text{\textit{ze1}} &= \begin{cases} 
\text{a. it is not more than this} \\
\text{b. you think this much/many is bad} \\
\text{c. I want you to know this much/many is not bad} 
\end{cases}
\end{align}

These definitions are worded with “semantic primes” based on the natural semantic metalanguage theory (for details of this theory, see [7], [8], and [16]). The definitions in (2) and (3) indicate that the meaning of \textit{zaa3} is entailed in the meaning of \textit{ze1}, i.e., line a. of (3), which can be seen as a paraphrase of “only.” The antecedent of “this” in (2) and (3a) is contained within in the proposition to which \textit{zaa3} or \textit{ze1} is attached. It is some quantifiable element plus any modifier(s), which may be a degree adverb or a numeral (e.g., “very little” or “three times”).

The antecedent of “this much/many” in lines (3b-c) is different from the antecedent of “this” in (3a). “This much/many” refers to a discourse element that the speaker assumes the listener considers bad, but which the speaker does not consider to be bad. The antecedent of “this” may or may not be the same as the antecedent of “this much/many.” For example, in (1b) the antecedent of “this” is “ten dollars,” but the antecedent of “this much/many” is a non-specific amount of money that the speaker understands pragmatically from the listener’s question. If, on the other hand, the listener’s question had been “Why are you wasting ten dollars on that?” then the antecedents of “this” in (3a) and “this much/many” in (3b-c) would both be “ten dollars,” but the former would refer to the “ten dollars” stated in the speaker’s sentence, and the latter to the “ten dollars” stated in the listener’s question.

It must be pointed out that the semantics of \textit{ze1} is more complicated than what is shown in (3), because \textit{ze1} does not always include the meaning of “only” and does not always include the meaning of “downplay” (See [6], chapter 2 for a detailed discussion). To account for this, the present study assumes that \textit{ze1} is polysemous. Its other meanings and any possible English equivalents of those other meanings are beyond the scope of this study, which relates only to the meaning of \textit{ze1} as defined in (3).

3. Methodology

Wakefield [15] carried out a preliminary study on the English equivalents of \textit{zaa3} and \textit{ze1}. Using the Hong Kong Cantonese Corpus (HKCanCor), which is a corpus of naturally occurring Cantonese speech, dialogues that included \textit{zaa3} and \textit{ze1} suffixed sentences were searched for and extracted. Native-bilingual participants then translated the target sentences of these dialogues into English. Wakefield concluded that \textit{zaa3} has an English equivalent in the form of the word “only” plus focus stress on the word that is put into focus. In contrast, no strong claim could be made about the English equivalent of \textit{ze1}, because it did not consistently translate into English as a specific pitch contour. It was proposed that the inconsistency in the translations was related to \textit{ze1} being polysemous. Two of the fifteen translations used a rise-fall-rise pitch, and it was suggested, based on Wakefield’s intuition, that this form of pitch has the meaning of “downplay” which has been ascribed to \textit{ze1} in the literature.

The purpose of the present study is to test the validity of Wakefield’s [15] tentative conclusions. The pitch contours that he proposed to be the equivalents of \textit{zaa3} and \textit{ze1} can be shown informally using the examples from (1) as follows (a. and b. are translations of \textit{zaa3} and \textit{ze1}, respectively):

\begin{align}
(1') \quad \text{a. “It’s only ten dollars.”} \\
\text{b. “It’s only ten dollars.”}
\end{align}

An actual translation from one of the native-bilinguals was as follows ([15], p. 250). The translated sentence was “It was only two.” The diameter of the circle in the figure represents the duration of the syllable “two,” which is the syllable that the “downplaying” tone is associated with:

![Figure 1: translation of \textit{ze1} on “It was only two”](image)

The meaning of this English contour is obviously not defined by (3) because it does not include the meaning “only” as represented in (3a). However, if the meaning of \textit{ze1} (or something very close to this meaning) is expressed in English as this particular form of intonation plus “just” or “only,” then (3) can be used as a guide for defining this pitch contour. The following procedures were used to verify whether \textit{ze1} has a (near) equivalent in English whose form includes a rise-fall-rise pitch contour.

3.1. Participants

Three Cantonese/English native bilinguals (one male; two female) provided three Cantonese-to-English translations of \textit{ze1}-suffixed sentences. Recordings of the participants speaking English and Cantonese were played to 10 native-English and 10 native-Cantonese speakers, respectively. All three were judged to be native Cantonese speakers by all ten of the native-Cantonese speakers. The male was judged to be a native English speaker by all of the native-English speakers, one female (female-a) was judged to be a native-English speaker by nine of the native-English speakers, and the other female (female-b) was judged to be a native-English speaker by seven of the native-English speakers. Perhaps female-b’s accent influenced some of the native-English speakers’ judgments of her English, because she spoke what sounded to the author to be a non-standard variety of British English, and those who judged her as less than native were speakers of Standard American English. Regardless of this, female-b was still judged to be a native-English speaker by the majority of native-English speakers who were asked to judge, and she is therefore very native-like if not native.
3.2. Translations

Three dialogues were constructed with the aim of eliciting translations of the meaning of ze1 as defined in (3), rather than one of its other meanings. The ze1-suffixed sentence was placed in a context that was intended to elicit the meaning of “downplay” from the native-bilingual participants. The three dialogues are shown in (4), (5) and (6) below. In each dialogue the target sentence (i.e., the ze1-suffixed sentence) is speaker B’s response to speaker A:

(4) A: Lei5 baai2 gam3 do1 ho2lok6 hai2dou6, dim2 you put so much coke here how gau3 wai2 baai2 kei4taa1 je5 aa3? enough space put other thing SFP “If you put that much coke here, how will there be enough space to put other stuff?”
B: Hai6 dak1 saam1gun3 ze1 be only three-CL ZE “There’re only three cans.”

(5) A: Waa4! Lei5 tai2 jat1tou3 hei3 ho2jii5 wow you watch one-CL movie can tai2 gam3 do1ci3 ge2, watch so many-CL(time) SFP “Wow! You can watch the same movie so many times!”
B: Ngo5 tai2-gwo3 loeng5ci3 ze1. I watch-EXP two-CL(time) ZE “I’ve only seen it twice.”

(6) A: Waa4! Lei5 jam2-zo2 gam3 do1 aa4? wow you drink-PERF so much PRT “Wow! You’ve drunk so much!”
B: Ngo5 jam2-zo2 hou2 siu2 ze1 I drink-PERF very little ZE “I only drank a little!”

Speaker A’s utterances all imply that s/he thinks the amount of the thing being referred to is excessive. It was predicted that this would cause the native-bilingual translators to downplay the amount, i.e., the number of coke cans, the number of movie viewings and the amount that was drunk.

Participants translated in separate sessions from each other and none heard the translations of the others.

3.3. F0 graphs

The pitch patterns are displayed within a two-octave range from the bottom to the top of each figure. The lines above and below a speaker’s median range represent a one-octave range. The center line marked with its hertz value represents the speaker’s estimated median pitch, which was calculated as the geometric mean of the speaker’s highest and lowest F0 point from within all of his or her own translations. This follows the method proposed by De Looze and Hirst [5].

4. Results

In Wakefield [15], 2 out of 15 translations of ze1 used a rise-fall-rise contour, one of which is shown in figure 1 above. The two tokens of the rise-fall-rise contour came from two different participants, meaning that none of the participants produced more than one token of this pitch contour. In this follow-up study, at least 5 of the 9 translations of ze1 used a rise-fall-rise contour; the male participant used it in all three of his translations; female-b used it in at least two of her translations (see discussion below); and female-a did not use it in any of her translations.

The three translations from the male participant are shown here in figures 2 to 4:

Figure 2: male’s translation of ze1 on “There’re only three cans”

The diameter of the circle shown in each figure represents the duration of the syllable(s) with which the “downplaying” tone is associated. In figure 2, the circle represents the duration of the syllable “cans,” and in figure 3, the syllable “twice.”

Figure 3: male’s translation of ze1 on “I’ve only seen it twice”

In figure 4 the diameter is the duration of the two syllables of the word “little.”:

Figure 4: male’s translation of ze1 on “I only drank a little”
There is clearly a rise-fall-rise shape inside the circle of each of the male’s translations.

Female-b’s first translation does not clearly show a rise-fall-rise F0 contour and it sounds to the author to be a muted version of this form of pitch. The circle shows the duration of “cans.”

None of female-a’s translations used a rise-fall-rise contour. Possible reasons for this are discussed below.

5. Discussion and Conclusions

The translations from female-a did not use the same pitch contour as the other two participants but her intonation was not neutral. She used what sounded to the author to be a defensive intonation, as if to indicate that she had done nothing wrong and didn’t like the fact that the conversational participant implied that she had. No attempt will be made here to further describe the meaning of female-a’s intonational forms other than to say that they sounded suitable to the contexts of the dialogues. There are different possibilities as to why she used a different form of intonation from the other two participants: she may express the meaning of “downplay” in a different form, or she may have added something additional to the discourse context in her mind which is suitable to another meaning of polysemous ze1. In other words: 1) it could be a speaker individual difference in relation to the intonational form of this meaning, or 2) it could be that the constructed dialogues did not work to elicit the meaning of ze1—as defined in (3)—from female-a, while these dialogues did work to elicit this meaning from the other two participants.

Hirst [10] said, “It is a remarkable thing in itself that many (though not all) untrained speakers are capable of reproducing the intonation of a sentence on a meaningless sequence of syllables” (p. 97, footnote). If mimicking the intonation of an utterance without changing languages is remarkable, then translating a semantically abstract segmental particle’s meaning from one language into another that expresses that meaning as a form of intonation is beyond remarkable. There are all sorts of things that could interfere with the cognitive processes involved. It is therefore very interesting that two participants used the same form of intonation without hearing each other’s translations. This is taken as a form of empirical evidence that a portion of the meaning of ze1 is (closely) equivalent to the meaning expressed by a rise-fall-rise pitch contour.

Based on the evidence of the translations, combined with the author’s native English-speaker intuition about this intonational form, it is proposed that the grammar of English includes this rise-fall-rise pitch contour, and that it functions to express the meaning that is paraphrased in lines b. and c. of the definition in (3), i.e., the meaning of ze1 less the meaning of “only” in (3a). Its meaning is proposed as follows:

\[(7) \Rightarrow \quad \text{a. I want you to think about this} \]
\[\text{b. you think this much/many is bad} \]
\[\text{c. I want you to know this much/many is not bad} \]

This tone is considered to be like an SFP in the sense that it functions to link the sentence to the discourse. As such its definition should represent this link by including deictic elements that refer to both the sentence (or a portion of it) and to something in the discourse. The antecedent of “this” in line (7a) is the proposition that the tone is associated with, and the antecedent of “this much/many” in (7b-c) is something in the discourse. The similarity between (3) and (7) is clear when you consider that a ze1 suffixed sentence such as “It’s ten dollars” translates into English as “It’s only ten dollars” plus a rise-fall-rise tone. The two definitions only differ by the first line. For (3a), the antecedent of “this” is the NP “ten dollars,” so it means: “it is not more than this: ten dollars.” For (7a), in contrast, the antecedent of “this” is the entire proposition, so it
means “I want you to think about this: It’s only ten dollars.” The two meanings are thus very similar.

Now that this English pitch contour has been “discovered” through translation from Cantonese, it can be examined independently of the SFP ze1. Unlike ze1, it does not entail the meaning “only” and should therefore be attachable to propositions which do not have this meaning. Consider the following constructed sentence, which seems to the author to be appropriate for this form of intonation. It is said by a mother to her son, who has just complained about having to look after his younger brother for an hour. The mother wants the older sibling to know that taking care of his younger brother for an hour is not as bad as he is implying.

(8) a. “It’s not gonna kill you.”
   b. “It’s only for an hour.”

The rise-fall-rise contour sounds natural on both “kill you” and “hour.” Only (8b) includes the adverb “only,” but the definition in (7) seems appropriate for (8a) as well, which can be seen more clearly if we restate these sentences based on (7):

(8a’) I want you to think about this (This much babysitting is not gonna kill you) you think this much [the amount of babysitting] is bad
I want you to know this much isn’t bad

(8b’) I want you to think about this (The babysitting is only for an hour) you think this much [the amount of babysitting] is bad
I want you to know this much isn’t bad

In both (8a’) and (8b’) the antecedent of “this much” is the same (i.e., “the amount of babysitting”) because it is a discourse element, and both sentences are part of the same discourse context.

It is worth noting that the “downplay” entailed in the meaning of ze1, which refers to some discourse element, will always coincide with a downplaying of some amount referred to in the speaker’s sentence. This is because ze1 entails the meaning “only.” The English rise-fall-rise contour also downplays the amount of some discourse element, but in contrast to ze1, this does not necessarily correspond to the downplaying of an amount in the speaker’s sentence, as was seen in (8a). The downplaying of the discourse element may even be accomplished by up-playing the amount of something that is referred to in the speaker’s sentence. Consider the examples in (9):

(9) [Two responses to “You haven’t studied enough.”]
   a. I studied for three hours.
   b. I did more than 20 exercises.

According to the author’s intuition, the rise-fall-rise tone sounds natural and appropriate when used on the word “hours” or “exercises” in a. and b., respectively. The speaker is downplaying the listener’s claim that not enough studying has been done. To accomplish this, the sentences in (9a-b) up-play the amount of studying done, which downplays the amount/degree of “insufficiency” of the amount of studying—or perhaps it can be seen as a downplay of the validity of the accusation. This can be seen more clearly by restating (9a-b) based on (7):

(9a’) I want you to think about this (I studied for 3 hours) you think this much [amount of studying] is bad
I want you to know this much isn’t bad

(9b’) I want you to think about this (I did more than 20 exercises)
you think this much [amount of studying] is bad
I want you to know this much isn’t bad

The antecedent of “this much” is the amount of studying that the speaker has done. The listener’s statement “You haven’t studied enough” implied that s/he thought the amount was insufficient and was therefore a “bad” amount. The speaker downplays the validity of the listener’s judgment by offering evidence which up-plays the amount of studying done, shown in the first lines of (9a’) and (9b’).

It is proposed that the form of intonation under discussion here exists as a floating tone morpheme in the English lexicon. Perhaps the definition in (7) could be refined, but if it is accepted that this pitch contour has some consistent meaning that is linked to its form, irrespective of the context, then it fits the standard definition of a morpheme. The data shows that it can be associated with a single syllable as in figures 1, 2, 3, 5, and 6, with two syllables as in figures 4 and 7, or even with four syllables according to the author’s intuition for (9b).

The observed pitch pattern is assumed to be the surface reflex of the interaction of this underlying floating tone morpheme with other accentual and boundary phenomena. The tone’s shape will vary from one occurrence to the next based on such interactions, but it is tentatively proposed that the rise-fall-rise pattern is a single tone rather than a combination of two or more. The initial rise sounds meaningful, and is therefore not considered to be merely a product of assimilation that precedes a fall-rise tone (a fall-rise tone seems to have a different meaning). Similarly, the second rise is not considered to be merely a high boundary tone that follows a rise-fall.

It is beyond the scope of this study to propose the precise phonological properties of this floating tone, and it is suggested that further research—using more data involving native English-speaker judgments and elicitations—be conducted with this goal in mind.

6. Acknowledgements

I am grateful to the native-bilingual participants who volunteered their time and effort to provide the translations for this study. I also wish to thank Daniel Hirst for his helpful comments. I thank the anonymous reviewers for their critical comments, and especially the reviewer who caused me to modify my definition in (7) and suggested a dialogue on which (9) is based. Any faults that remain in the paper are of course my own.

7. References


12. C. Leung, *當代香港粵語語助詞的研究 [A Study of the Utterance Particles in Cantonese as Spoken in Hong Kong]*. Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong, 2005.


