



Research Article

Temporal coordination between focus prosody and pointing gestures in Cantonese [☆]



Holly Sze Ho Fung, Peggy Pik Ki Mok ^{*}

Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Shatin, Hong Kong

ARTICLE INFO

Article history:

Received 11 July 2017

Received in revised form 19 July 2018

Accepted 23 July 2018

Keywords:

Co-speech gesture

Focus prosody

Cantonese

Multimodal communication

ABSTRACT

This study investigates the temporal relationship between focus prosody and co-speech pointing gestures in Hong Kong Cantonese. Previous studies have generally shown a close temporal proximity between prosodic and gestural prominence: Gestural prominence tends to be aligned with stressed syllables or words. However, this finding was based solely on studies of stress and pitch-accent languages, and no study has yet tested the phenomenon in a non-stress tone language. Ten native speakers of Hong Kong Cantonese participated in a picture-verification task in which pointing was elicited along with verbal corrections. The acoustic results showed that the corrective focus was marked solely by an on-focus durational increase. The gestural results revealed that there was an alignment between prosodic and gestural prominence, as most of the gesture apices were produced within the focused words. However, in contrast to previous findings, no significant effect of F0 (tone) or focus position was found. Instead, most speakers consistently aligned their apices with the same syllable position in disyllables. Based on the current findings, the prosodic anchor of prosody-gesture alignment is suggested to be the focused word in this language.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. The phonological synchrony rule

Speech communication is essentially multimodal. Apart from spoken words, information is also conveyed by co-speech gestures, such as bodily movements that one produces as one speaks, which include both manual movements (for example, pointing and hand beats) and non-manual movements (such as head movements).

Despite the formal differences between speech and gesture, McNeill (1992, 2005), among others (Goldin-Meadow, 1998; Kendon, 1972, 2000, 2004; Kelly, Manning, & Rodak, 2008), have argued that gesture and language comprise one system. He gave five reasons for his argument:

- (1) Gestures occur almost exclusively during speech;

- (2) both speech and gestures convey similar if not the same semantic meanings, in addition to serving identical pragmatic functions;
- (3) the two modalities exhibit temporal synchrony;
- (4) the development of gestures in children mirrors that of language, both beginning with concrete deixis and ending on the discourse level; and
- (5) impaired speech and gestures in aphasic patients show parallel patterns, as both lack coherence but preserve meaning in Broca's aphasics, and exhibit fluency without interpretability in Wernicke's ones.

All the evidence shows not only similarities between speech and gestures on the surface level, but also suggests a shared underlying system processing both of them.

With regard to their regular resemblance in semantic meaning, pragmatic functions and timing (reasons (2) and (3) above), McNeill (1992) proposed three synchrony rules for speech and gesture, namely the semantic, pragmatic and phonological synchrony rules. The first two state that co-occurring speech and gesture present and perform the same meaning and pragmatic functions, whereas the last, phonological synchrony, which is the most relevant to the present study, states that “the stroke of the gesture precedes or ends at, but

[☆] This work was supported by the Graduate School of the Chinese University of Hong Kong awarded to the first author (postgraduate studentship, 2013–2015).

^{*} Corresponding author.

E-mail addresses: hollyfung@link.cuhk.edu.hk (H.S.H. Fung), peggymok@cuhk.edu.hk (P.P.K. Mok).

does not follow, the phonological peak syllable of speech” (p. 26). The stroke here refers to the only obligatory and the most prominent part of a gesture, preceded and followed optionally by the preparation and recovery phases (Kendon, 1972, 1980; Kita, 1990; McNeill, 1992) (although the apex of the stroke has been considered the unit of gestural prominence instead in many recent studies, as will be reviewed in Section 1.2). In other words, what the rule suggests is, when speech and accompanying gestures unfold, the most prominent parts of both channels are linked temporally.

1.2. Previous studies on the temporal relationship between prosodic and gestural prominence

A number of production studies have been conducted to investigate the occurrence of gestural prominence relative to prosodic prominence in simultaneous speech in different intonation and pitch-accented languages. They can be categorised into four groups according to whether or not an alignment between prominent units in speech and gesture was found and, if it was, if the prominent unit in speech, or the *prosodic anchor*, was (a) a stressed/accented word, (b) a stressed/accented syllable, or (c) an F0 peak. The four groups of studies are reviewed as follows.

A few studies found no effect of the change in lexical/nuclear stress position on the timing of gestural prominence, usually measured by the apex (in other words, the maximal displacement of the gesturing hand/body part). For example, De Ruiter (1998, Experiment 1) found that a change in the metrical structures (stress-initial versus stress-final) of nouns elicited in definite *determiner + noun* responses in Dutch had no significant effect on the apex times of accompanying pointing gestures, although apices did occur before accented syllables as predicted by the phonological synchrony rule. Furthermore, in a picture-naming task, Rusiewicz, Shaiman, Iverson and Szuminsky (2013) elicited pointing gestures co-produced with American English sentences, in which contrastive stress was placed on either the first or the second syllables of the target words, which were dimorphemic, trochaic compound nouns. In line with De Ruiter’s (1998) finding, the results showed no significant effect of contrastive stress position on the timing of the gesture apex.

Nonetheless, more studies have provided evidence for close temporal alignment or covariation between prosodic and gestural prominences, although with different suggestions regarding the prosodic anchor of alignment. Some have suggested that it is the stressed/focused word. For example, Roustan and Dohen (2010) elicited contrastive focus on either the subject or the object (both being CVCV words) in simple SVO sentences in French, which were accompanied by pointing, beat or control (in other words, button pressing) gestures. They found that the apices of the pointing gestures were consistently aligned with the articulatory target of one of the vowels of the focused word.

There is also evidence for the prosodic anchor being the pitch-accented/stressed syllable rather than the word carrying that syllable. Following the experimental settings of his first experiment, De Ruiter (1998, Experiment 2) elicited pointing gestures co-occurring with Dutch noun phrases in the structure *definite determiner + colour adjective + noun*, in which

contrastive stress was placed on four possible positions (two content words \times two metrical structures). The results showed a significant positive correlation between the beginning of an apex and the onset of a stressed syllable. By contrast, alignment between the gesture apex and stressed syllable is achieved differently in Brazilian Portuguese. In a study by Rochet-Capellan, Laboissière, Galván and Schwartz (2008), speakers of Brazilian Portuguese were asked to identify and point at pictorial targets, which had either trochaic or iambic labels. The results showed that the stressed syllables of both trochaic and iambic words were in sync with gesture apices (that is, the period of time during which the finger remained pointing at the pictorial target), but in different ways: Stressed syllables of the trochees were aligned with the *beginning* of apices, whereas those of the iambs were aligned with the *end* of them. These experimental results were in line with observations from naturalistic studies. Examining manual gestures produced by American English speakers during natural conversations, Loehr (2012) found that pitch accents were only +17 milliseconds ahead of the nearest gesture apices on average. Also studying spontaneous gestures accompanying English speech, Jannedy and Mendoza-Denton (2005) found that 95.7% of all the observed apices co-occurred with a pitch accent. Evidence of a close alignment is also provided by studies on non-manual gestures. For example, Esteve-Gibert, Borràs-Comes, Swerts and Prieto (2014) investigated head movements produced by Catalan speakers in a semi-spontaneous setting in which target words with different metrical patterns were elicited naturalistically, and found that the apices of the head gestures were aligned with accented syllables. Similarly, Ambrazaitis, Lundmark and House (2015) and Ambrazaitis and House (2017) found that head beats and eyebrow movements were closely associated with focal pitch accents in Swedish broadcast news.

Yet another view is that it is the F0 peak of the pitch-accented/stressed syllable that attracts gestural prominence. Leonard and Cummins (2011) studied elicited beat gestures co-occurring with English sentences and found that, among three different possible speech landmarks of speech-gesture alignment, including the rhythmic pulse (P-centre), the vowel onset and the F0 peak of the accented syllable, the gesture apex was aligned closest to the F0 peak. Similarly, in a controlled setting, in which corrective focus structures in Catalan were elicited simultaneously with pointing gestures, Esteve-Gibert and Prieto (2013) found that the correlation between the gesture apex and the F0 peak was the strongest when compared to other pairs of speech and gestural prominent units, including

- (1) the apex and the end of the accented syllable,
- (2) the stroke onset and the F0 peak, and
- (3) the stroke offset and the F0 peak.

As reviewed above, a number of prosodic units have been proposed as the prosodic anchor of speech-gesture coordination. They include (from larger to smaller)

- (1) the focused/accented word in a sentence,
- (2) the stressed/accented syllable of that word, and
- (3) the F0 peak of that syllable.

Download English Version:

<https://daneshyari.com/en/article/11005573>

Download Persian Version:

<https://daneshyari.com/article/11005573>

[Daneshyari.com](https://daneshyari.com)