摘要

副语言信息指说话者加在自己当下所说内容之上、利用各种语音线索（如音高、时长等）去表达与交际意图相关的信息。根据以往研究，有的副语言信息反映说话者与听话者之间的人际关系（对人），如支配性语音；有的副语言信息反映说话者对自己当下所说内容的立场、态度或意愿（对事），如确信性语音。交际意图指向可以分为对人、对事两种，但听话者是否可以真正利用语音线索区分对人、对事这两类态度性语音，迄今证据不足。另一方面，已有有关于副语言信息（态度性语音）的语音学、心理语言学以及认知神经科学的实证研究仅聚焦于态度性语音传递的具体交际意图（如确信、支配等），但对态度性语音如何传递的交际意图指向（对人、对事），学界鲜有探究。

本文利用行为实验、语音分析与 ERP 实验，探究听话者是否可以利用副语言信息区分说话者的交际意图指向（对人 vs. 对事）。

研究一在行为感知层面上探究听话者在利用副语言信息区分说话者的交际意图指向（对人、对事）时加工模式是否具有差异。实验操纵了交际意图指向（对人、对事、中立）与句子的呈现方式（关键句、引导语加关键句；引导语中表达态度的词汇语义信息能帮助判断说话者的交际意图指向，如“我十分确信”；关键句中仅有语音信息区分说话者的交际意图指向）。被试听到说话者的言语，通过按键判断说话者的交际意图指向（机会概率为 33%）。结果表明：对关键句条件的判断正确率（55%）显著小于对引导语加关键句条件的判断正确率（97%）。当仅呈现关键句时，对人判断的正确率（24%）显著小于对事的判断（86%）；引导语的出现使得对人（95%）、对事（99%）判断正确率的差异消失。说明嗓音副语言在编码对人、对事的交际意图指向的信息时可能存在差异，基于嗓音副语言识别交际意图指向对词汇语义信息可能存在不同程度的依赖。

研究二在声学、语音层面上探究听话者如何利用副语言信息区分说话者的交际意图指向（对人、对事）。为了保证本研究语料的有效性与各条件语料的充足性，筛选研究一中感知判断正确率高于 55%（机会概率为 33%）的呈现方式为关键句的语料（对事 1170 句，对人 163 句，中立 164 句）进入声学分析。对不同交际意图指向语句的 11 个声学参数（平均基频、基频范围、平均音强、音强范围、平均谐噪比、谐噪比的标准差、最大基频时间、最小基频时间、最大音强时间、最小音强时间、句子时长）进行了语音分析，结果表明：在这 11 个声学参数上，对事的各种数值均显著大于对人的各种数值，说明听话者能基于多个语音线索区分说话者的交际意图指向。

研究三在神经认知层面上探究听话者如何利用副语言信息区分说话者的交

际意图指向（对人、对事）。为了保证本研究语料的有效性与各条件语料的充足性，筛选研究一中感知判断正确率不低于 75%（超过机会概率 2 倍）的呈现方式为引导语加关键句的材料（共 80 套）进入 ERP 实验。实验操纵了交际意图指向（对人 vs. 对事）和句子中引导语中词汇语义所表达的态度与关键句声学、语音学特征所表达的态度之间的一致性（一致 vs. 不一致），被试通过按键判断关键句的交际意图指向。行为结果表明，对人判断的正确率（93%）相比对事判断的正确率（99%）更低，对人判断的反应时（1601 ms）相比对事的反应时（1403 ms）更慢。仅在对人判断时，对不一致条件句判断的正确率（90%）比对一致条件句判断的正确率更低（95%）、对不一致条件句判断的反应时（1738 ms）比对一致条件句判断的反应时（1470 ms）更低更慢。事件相关电位（ERP）结果表明：在关键句上，N1（160 - 200 ms）在对事条件（-0.485 μV）下比在对人条件（-0.331 μV）下有更大的负向波幅，且仅在对人条件下，不一致条件（-0.410 μV）比一致条件（-0.252 μV）有更大的负向波幅；P2（250 - 290 ms），在对人条件下（1.044 μV）比在对事条件（0.974 μV）下有更大的正向波幅，且仅在对人条件下，一致条件（1.283 μV）比不一致条件（0.806 μV）有更大的正向波幅；晚期持续正波（900 - 1600 ms）在对人条件（-0.512 μV）下比在对事条件（-0.739 μV）下有更大的正向波幅，且仅在对人条件下，不一致条件（-0.423 μV）比一致条件（-0.600 μV）有更大的正向波幅。这些结果表明嗓音副语言信息编码的不同交际意图指向（对人、对事）在声学解码、注意资源分配与语用推理整合等不同嗓音表情解码过程上都表现出差异。这些嗓音解码的神经认知过程也在意图指向发生变化时表现出差异，且仅当说话人的嗓音编码对人的交际意图指向时才表现出差异。

三个研究的结果揭示了对嗓音副语言信息编码的不同交际意图指向（对人、对事）的解码在行为、声学与神经认知层面上具有差异，表明听话者能够利用副语言信息区分说话者的交际意图指向。这些发现扩展了嗓音表情解码的认知加工模型。

关键词：副语言信息；态度性语音；交际意图；嗓音表情；事件相关电位

Abstract

Paralinguistic information refers to the information added to the utterance by the speaker who employs various acoustic cues (e.g. pitch and duration) to express communicative intention. According to the previous studies, some paralinguistic information reflect the interpersonal relationship between the speaker and the listener (listener-directed), such as dominant speech; some paralinguistic information reflect the speaker's stance, attitude or will toward the content being uttered by himself or herself (proposition-directed), such as certain speech. Direction of communicative intenion can be classified into two categories (listener-directed and proposition-directed). However, whether listeners can truly employ acoustic cues to differentiate these two categories of attitudinal speech has been scantly evidenced so far. On the other hand, there exists research of phonetics, psycholinguistics and cognitive neurosciences on paralinguistic information (attitudinal speech) merely focusing on specific communicative intention (e.g. certainty and dominance) expressed by attitudinal speech. However, how attitudinal speech convey direction of communicative intention (listener-directed / proposition-directed) has been rarely explored in the field.

This thesis employed a behaviorial experiment, an acoustic analysis and an ERP experiment to investigate whether listeners can employ paralinguistic information to differentiate speakers' direction of communicative intention (listener-directed vs. proposition-directed).

Study One investigates whether listeners' processing patterns of employing paralinguistic information to differentiate speakers' direction of communicative intention (listener-directed, proposition-directed) differ at the behavioral level. The experiment manipulated direction of communicative intention (Listener-Directed, Propositional-Directed and Neutral) and modes of presentation (Target Sentence and Leading Phrase plus Target Sentence; the lexical-meaning information in the leading phrases (for example, "I'm pretty sure.") can help judge the speakers' direction of communicative intention; merely acoustic information can be employed to differentiate the speakers' direction of communicative intention in the target sentences. Participants listened to the speakers' utterance and judged the speakers' direction of communicative intention by pressing a button (the chance level: 33%). Results showed that Target Sentence (55%) was significantly less accurately judged than Leading Phrase plus Target Sentence (97%). When merely a target sentence was presented, Listener-

Directed (24%) was significantly less accurately judged than Proposition-Directed (86%); emergence of a leading phrase made the differences between the accuracy of Listener-Directed (95%) and the accuracy of Proposition-Directed (99%) disappear. These suggest that vocal paralinguistic information may differ during the encoding of direction of communicative intention (listener-directed and proposition-directed) and that lexical-meaning information is relied on to different extents during the recognition of direction of communicative intention based upon vocal paralanguage.

Study Two investigates how listeners employ paralinguistic information to differentiate speakers' direction of communicative intention (listener-directed and proposition-directed) at the acoustic and phonetic levels. To ensure the validity of the materials in this study and the sufficiency of the materials in various conditions, the materials whose mode of presentation is Target Sentence with the accuracy of perceptual judgement higher than 55% (the chance level: 33%) entered the acoustic analysis (1170 proposition-directed, 163 listener-directed, 164 neutral). Analysis of 11 acoustic parameters (mean f0, f0 range, mean intensity, intensity range, mean HNR, standard deviation of HNR, time for maximal f0, time for minimal f0, time for maximal intensity, time for minimal intensity, utterance duration) to the different direction of communicative intention was conducted. Results showed that all the values of the 11 acoustic parameters of Proposition-Directed were higher than those of Listener-Directed, suggesting that listeners can differentiate speaker's direction of communicative intention according to various acoustic cues.

Study Three investigates how listeners employ paralinguistic information to differentiate speakers' direction of communicative intention (listener-directed / proposition-directed) at the neurocognitive level. To ensure the validity of the materials in this study and the sufficiency of the materials in various conditions, the materials whose mode of presentation is Leading Pharse plus Target Sentence with the accuracy of perceptual judgement no lower than 75% (above twice the chance level) entered the ERP experiment (80 sets of materials). The experiment manipulated direction of communicative intention (Listener-Directed vs. Proposition-Directed) and contingency between the attitude expressed by the lexical meaning of the leading phrases and the attitude expressed by the acoustic and phonetic features of the target sentences (Contingent vs. Incontingent). Participants judged the direction of communicative intention expressed in the target sentences by pressing a button. Behavioral results showed that Listener-Directed (93%) was less accurately judged than Proposition-

Directed (99%). Listener-Directed (1601 ms) was more slowly judged than Proposition-Directed (1403 ms). Incontingent (90%) was less accurately judged than Contingent (95%) and Incontingent (1738 ms) was more slowly judged than Contingent (1470 ms) merely when Listener-Directed was judged. Results of event-related potentials (ERPs) showed: in the target sentences, Proposition-Directed (-0.485 μV) elited a stronger N1 (160 – 200 ms) than Listener-Directed (-0.331 μV); Incontingent (-0.410 μV) elicited a stronger N1 than Contingent (-0.252 μV) only in Listener-Directed. Listener-Directed (1.044 μV) elited a stronger P2 (250 – 290 ms) than Proposition-Directed (0.974 μV); Contingent (1.283 μV) elicited a strong P2 than Incontingent (0.806 μV) only in Listener-Directed. Proposition-Directed (-0.512 μV) elited a stronger late sustained positivity (900 – 1600 ms) than Listener-Directed (-0.739 μV); Incontingent (-0.423 μV) elicited a stronger late sustained positivity than Contingent only in Listener-Directed (-0.600 μV). These results suggest that different direction of communicative intention (listener-directed, proposition-directed) encoded by vocal paralinguistic information differ during different processes of vocal expression decoding, such as acoustic decoding, attentional allocation and pragmatic inference and integration. These neurocognitive processes of vocal expression decoding also differ when direction of communicative intention is changed. The differences emerge merely when speakers encode listener-directed direction of communicative intention.

Results of the three studies reveal that the decoding of different direction of communicative intention (listener-directed, proposition-directed) encoded by vocal paralinguistic information differ at the behavioral, acoustic and neuroconitive levels, suggesting that listeners can employ paralinguistic information to differentiate speakers' direction of communicative intention. These findings extend the cognitive processing model of vocal expressions.

Keywords: paralinguistic information; attitudinal speech; communicative intention; vocal expression; ERPs