The De Se Theory of Indexicals: Convergent Evidence from Philosophy, Typology, Developmental Psychology, and Linguistic Semantics

This talk proposes a De Se Theory of Indexicals, wherein 1\textsuperscript{st} and 2\textsuperscript{nd} person pronouns indicate reference de se (= self-ascription). Self-ascription is a property long observed for 1\textsuperscript{st} person pronouns (Castañeda 1968; Perry 1979; Kaplan 1977) and extended here to 2\textsuperscript{nd} person as well. Prative 1\textsuperscript{st} and 2\textsuperscript{nd} person pronoun features \([spk]\) and \([addr]\) designate the speaker and addressee, respectively, as the self-ascriber. Anyone who is not a designated self-ascriber for a given pronoun can only interpret it indirectly by inferring the self-ascriber’s interpretation, a process requiring Theory of Mind (ToM), i.e. the ability to impute mental states to others (Premack and Woodruff 1978). Only via features \([spk]\) or \([addr]\), directly or indirectly, can a pronoun be knowingly used to refer to a speaker or addressee (e.g. \textit{you.PL} is marked \([addr]\); it lacks \([spk]\), so its reference set can’t include the speaker). This De Se Theory is supported by convergent evidence from multiple domains:

(i) \textit{Philosophy}. It solves the Problem of the Essential Indexical (Perry 1979). Crimmins (1992:163-5) showed this already for 1\textsuperscript{st} person; here it is extended to 2\textsuperscript{nd} person.

(ii) \textit{Typology}. It explains the herefore mysterious absolute morphological universal (Greenberg 1988, Noyer 1997, Cysouw 2003, Bobaljik 2008): unlike the plural semantics of common nouns (the dogs: Every member of the reference set is a dog), 1pl and 2pl in all languages have ‘associative’ semantics (we/you.PL: Some member of the reference set is the speaker/addressee). See Table I. On the De Se Theory, the person feature specifies (\(\forall\)) quantification over speakers or addressees— not quantification over the members of the pronoun reference set— so the person feature cannot limit reference to ‘only addressees’ or ‘only speakers’.

(iii) \textit{Developmental psychology}. Two groups have a ToM deficit: children under about 3.5 y.o. (Sodian 2006, i.a.); and children with autism (Baron-Cohen, Leslie, and Frith 1985, i.a.). The De Se Theory explains their special problems with 1\textsuperscript{st}/2\textsuperscript{nd} person pronouns. Young children have the most difficulty with speaker-production of 2\textsuperscript{nd} person and addressee-comprehension of 1\textsuperscript{st} person pronouns (Charney 1980; Chiat 1986)— exactly the non-self-ascribed ones, requiring ToM on the De Se Theory. Children with autism frequently reverse 1\textsuperscript{st} and 2\textsuperscript{nd} person (13\% reversed in one study (Tager-Flusberg 1994, 184)).

(iv) \textit{Linguistic semantics}. Each addressee \(x\) hearing 1\textsubscript{a} interprets \textit{your} as referring to \(x\): if Tom and Mary are both addressees, Tom knows to write his own name, not Mary’s. Each addressee \(x\) hearing 1\textsubscript{b} interprets \textit{you} as referring to a set that includes \(x\) (e.g. ‘\(x\) and \(x\)’s wife’). The De Se Theory predicts this, since each addressee self-ascribes membership in the reference set (4, 5). On standard double-indexing theories \textit{you} is anchored to an addressee, wrongly allowing one addressee to interpret \textit{you} as referring to another. (One fix: Identify the addressee index with a variable over elements of the pronoun’s reference set and let a quantifier bind that variable from outside the illocutionary operator. But this is a stipulation.)

Our formalization builds on Crimmins and Perry (1989) and Crimmins (1992, 163-5). \textit{Beliefs} are private cognitive particulars, with structures built from \textit{notions} and \textit{ideas}. The content of an agent’s belief is a proposition (a public classification of circumstances in the world, built from \textit{individuals} and \textit{relations}); the content of an agent’s \textit{notions} and \textit{ideas} are \textit{individuals} and \textit{relations}, respectively. A basic relation \textit{Believes}(A,p,\(\pi\)) relates an agent \(A\), a proposition \(p\), and a belief structure \(\pi\). The grammar builds the agents’ belief structures, which are mapped to their contents. A language \(L\) is characterized by the partial function \(V: \langle X, \phi^{\hat{sl}} \rangle \mapsto B\), where \(X\) is the set of \(L\)-competent agents, \(\phi^{\hat{sl}}\) is a linguistic expression uttered by a set \(S\) of speakers (a singleton except in mass speaking) to a set \(H\) of addressees, and \(B\) ranges over constituents of the agent’s belief structures. The notions held by an agent \(x\) include a self-notion \(x^3_{n_{self}}\) that is necessarily a notion of \(x\) (4). Self-ascription is ascription via the self-notion. A partial function, \(V\) maps a 1\textsuperscript{st} (2\textsuperscript{nd}) person pronoun to a set containing the speaker’s (addressee’s) self-notion \(x^3_{n_{self}}\). \(V\) returns no value for other agents interpreting the pronoun, who must therefore induce its meaning via Theory of Mind. See 7 for a sample derivation.
1. a. Write your name at the top of the page.  
   b. How often do you kiss each other?  

Table I. Seven logically possible meta-persons; only four attested pronoun types (Bobaljik 2008)

<table>
<thead>
<tr>
<th>Possible</th>
<th>Attested</th>
<th>Privative features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+2: speaker(s) &amp; addressee(s) only</td>
<td>‘inclusive’</td>
<td>[spk, addr]</td>
</tr>
<tr>
<td>1+2+3: speaker(s), addressee(s) &amp; other(s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: speaker(s) only</td>
<td>‘exclusive’</td>
<td>[spk]</td>
</tr>
<tr>
<td>1+3: speaker(s) &amp; other(s) only</td>
<td></td>
<td>[addr]</td>
</tr>
<tr>
<td>2: addressee(s) only</td>
<td>‘second person’</td>
<td></td>
</tr>
<tr>
<td>2+3: addressee(s) and other(s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3: other(s) only</td>
<td>‘third person’</td>
<td></td>
</tr>
</tbody>
</table>

Suppose Mary and Paula, both speakers of English, know the city of Austin, Texas by the name Austin:
E.g. for Mary, the word Austin translates as Mary’s notion of Austin ("n_Austin").

2. English translation function V.
   a. V((Mary, [Austin]S,H)) = "n_Austin"  
   b. V((Paula, [Austin]S,H)) = "n_Austin"  
   c. V((Mary, [likes]) = \( \lambda y x \{ \lambda_i \text{likes}, x, y \} \)  
   d. V((Paula, [likes]) = \( \lambda y x \{ \lambda_i \text{likes}, x, y \} \)

(Attested) Function ContentOf

3. The function ContentOf
   a. ContentOf("n_Austin") = Austin  
   b. ContentOf("n_Austin") = Austin  
   c. ContentOf("likes") = Likes  
   d. ContentOf("likes") = Likes

(Contextual indices S and H: set of speakers and addressees).

4. Self-Notion Axiom. Necessarily, \( \forall x [\text{ContentOf}(n_{self}) = x] \).

5. Pronoun features: number; [spk] & [addr] (informal). A pronoun denotes a set G, where:
   a. [NUM {sg/pl/dual/...}]: G has one/more than one/two/... members.  
   b. [spk]: Every speaker self-ascribes membership in G.  
   c. [addr]: Every addressee self-ascribes membership in G.

6. Pronoun features [spk] and [addr] as constraints on V.
   a. \( \forall s \in S \left[ V(s, \{ \text{spk} \}, S^H) \supseteq \{ n_{self} \} \right] \); \( \forall x \in S \left[ V(x, \{ \text{spk} \}, S^H) \right. \) is undefined
   b. \( \forall h \in H \left[ V(h, \{ \text{addr} \}, S^H) \supseteq \{ n_{self} \} \right] \); \( \forall x \in H \left[ V(x, \{ \text{addr} \}, S^H) \right. \) is undefined

7. [ I like Austin ]_{(Mary), (Paula)} (Mary speaking to Paula)
   a. speaker: Believes(Mary, \( \langle \langle \text{Likes}; \text{Mary}, \text{Austin} \rangle \rangle \), \( \{ \lambda_i \text{likes}, n_{self}, n_{Austin} \} \))  
   b. addressee: Believes(Paula, \( \langle \langle \text{Likes}; \chi, \text{Austin} \rangle \rangle \), \( \{ \lambda_i \text{likes}, n_{Austi} \rangle \rangle n_{Austin} \})

A place-holder \( \eta \) shows where \( V \) returns no value (see 6a). Paula solves for \( \eta \) by building a model of Mary’s belief state using 6a: \( \eta = \text{Mary} \). She lets \( \chi = n_{Mary} \) since ContentOf(\( n_{Mary} \)) = Mary.

Selected references