Person-Voice Interactions in Optimality Theory

JUDITH AISSEN (1999)
JOAN BRESNAN, SHIPRA DINGARE, AND CHRIS MANNING (2001)
JUDITH AISSEN AND JOAN BRESNAN (2002)

[227B, Spring 2002–3]
1. Background: Markedness Hierarchies in Syntax

“... the most natural kind of transitive construction is one where the A is high in animacy and definiteness, and the P is lower in animacy and definiteness; and any deviation from this pattern leads to a more marked construction.”

— Comrie (1989: 128)

Subject > Object
High in animacy > Low in animacy
High in definiteness > Low in definiteness
Local person (1,2) > Third person
Crosslinguistically, the more marked transitive configurations may be flagged (for example, by inverse or case morphology) or avoided (for example, by passivization).

*The spread of markedness in transitive actives by person/role:*

<table>
<thead>
<tr>
<th>Agent ↓</th>
<th>Patient →</th>
<th>Local person</th>
<th>Third person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local person</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third person</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Lummi examples:\textsuperscript{a, b}

\[
\begin{align*}
\text{xči-t-s} & \quad \text{cē swəyʔqəʔ} & \quad \text{cē swiʔqoʔət} \\
\text{know-TR-3.ERG} & \quad \text{the man} & \quad \text{the boy} \\
\text{‘The man knows the boy’}
\end{align*}
\]

\[
\begin{align*}
\text{xči-t-ŋ} & \quad \text{cē swiʔqoʔət} & \quad \text{cē swəyʔqəʔ} \\
\text{know-TR-PASS} & \quad \text{the boy} & \quad \text{by the man} \\
\text{‘The boy is known by the man’}
\end{align*}
\]

* \text{‘The man knows me/you’}

\[
\begin{align*}
\text{xči-t-ŋ=sən/=sxʷ} & \quad \text{cē swəyʔqəʔ} \\
\text{know-TR-PASS=1/2.SG.NOM} & \quad \text{by the man} \\
\text{‘I am/you are known by the man’}
\end{align*}
\]

\textsuperscript{a}Lummi is a dialect of Coast Salish spoken in British Columbia. The suffix -TR is a verbaliser; -ERG denotes ergative, the case of third person transitive subjects.

\textsuperscript{b}Jelinek, Eloise and Richard Demers. 1983. ‘The agent hierarchy and voice in some Coast Salish languages,’ \textit{IJAL} 49, 165–85.

Picurís examples:\(^a,^b\)

\[S\œnene \ mœn-\text{-}q\text{n}.
\]
man \ see-PAST

‘The man saw him.’

\[Mœn-mia-\text{-}q\text{n} \ sœnene-pa.
\]
see-PASS-PAST \ man-OBL

‘He was seen by the man.’

\[^*\] \ ‘The man saw me.’

\[Ta-mœn-mia-\text{-}q\text{n} \ sœnene-pa.
\]
1SG \ SUBJ\text{\_intrans} \ see-PASS-PAST \ man-OBL

‘I was seen by the man.’

\(^a\) Picurís, unrelated to Lummi, is a dialect of Northern Tewa spoken in New Mexico. Third persons are zeros except as objects of a verb with local person subject.


Passivization in Lummi, Picurís:

<table>
<thead>
<tr>
<th>Agent ↓</th>
<th>Patient →</th>
<th>Local person</th>
<th>Third person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local person</td>
<td>no</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Third person</td>
<td>oblig</td>
<td></td>
<td>opt’l</td>
</tr>
</tbody>
</table>

*The classic view of markedness* captures implicational universals (languages share unmarked constructions and differ with respect to marked ones), but it *does not explain how the language-particular manifestations of the hierarchies can vary from language to language and even within the same language, while sharing universal trends in direction.*
Like Lummi and Picurís are Nootka (Southern Wakashan, British Columbia) (Klokeid 1969, Whistler 1985, Emanatian 1988) and Chamorro (Western Malayo-Polynesian, Guam and Northern Mariana Islands, Chung 1998, Cooreman 1987)—also unrelated.


<table>
<thead>
<tr>
<th>Passive:</th>
<th>Inverse:</th>
</tr>
</thead>
<tbody>
<tr>
<td>intransitive</td>
<td>transitive</td>
</tr>
<tr>
<td>patient Subject</td>
<td>patient Object</td>
</tr>
<tr>
<td>oblique case marking on agent</td>
<td>non-oblique agent</td>
</tr>
<tr>
<td>omissibility of indefinite agent</td>
<td>non-omissibility</td>
</tr>
</tbody>
</table>

Mithun (1999: 227) concludes of Picurís, “There is no question that these constructions are formally passive.”
2. Markedness Hierarchies in Optimality Theory

(Aissen 1999)\textsuperscript{a}

<table>
<thead>
<tr>
<th>A structural, binary scale</th>
<th>An \textit{n}-ary, substantive scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntactic structure:</td>
<td>Role scale:</td>
</tr>
<tr>
<td>Subject &gt; nonSubject</td>
<td>protoAgent &gt; protoPatient</td>
</tr>
<tr>
<td>Subject &gt; Object</td>
<td>Person scale:</td>
</tr>
<tr>
<td>Subject &gt; Oblique</td>
<td>Local (1,2) &gt; 3</td>
</tr>
<tr>
<td></td>
<td>Topicality scale:</td>
</tr>
<tr>
<td></td>
<td>Topical &gt; nonTopical</td>
</tr>
</tbody>
</table>

Harmonically Aligned scales:

\[ S_{ag} \succ S_{pt} \quad S_{1,2} \succ S_{3} \quad S_{\text{topical}} \succ S_{\text{nontopical}} \]
\[ O_{pt} \succ O_{ag} \quad O_{3} \succ O_{1,2} \quad O_{\text{nontopical}} \succ O_{\text{topical}} \]

\[ \ldots \]

\textsuperscript{a}Aissen, Judith. 1999. ‘Markedness and subject choice in optimality theory,’ \textit{Natural Language & Linguistic Theory} 17, 673–711.
Classic markedness theory: Disharmonic argument types are either avoided or flagged by formal marking.

Implementation in Optimality Theory (OT) as universal markedness constraint subhierarchies:

\[
* S_{pt} \gg * S_{ag} \quad * S_{3} \gg * S_{1,2} \quad * S_{nontopical} \gg * S_{topical}
\]
\[
* O_{ag} \gg * O_{pt} \quad * O_{1,2} \gg * O_{3} \quad * O_{topical} \gg * O_{nontopical}
\]
\[
\ldots
\]

Note 1: Other constraints may be interleaved, but the subhierarchy structure is preserved.

Note 2: Harmonic alignment captures markedness reversal—what’s unmarked for subjects is marked for objects, and conversely.
An Example: The Passive

Avoidance of disharmonic patient subjects (*S_{pt}) contributes to the crosslinguistic markedness of passives.

In OT, the universal subhierarchy *S_{pt} \gg *S_{ag} disfavors the passive compared to the active (all else being equal):

<table>
<thead>
<tr>
<th></th>
<th>*S_{pt}</th>
<th>*S_{ag}</th>
</tr>
</thead>
<tbody>
<tr>
<td>input: v(ag,pt)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>passive: S_{pt}, Obl_{ag}</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>active: S_{ag}, O_{pt}</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Then why is there passivization at all?
Other constraints favor the passive. \(^a\) English avoids nontopical subjects:

<table>
<thead>
<tr>
<th>Input: v.ag/new, pt</th>
<th>*S(_{\text{nontopical}})</th>
<th>*S(_{pt})</th>
<th>*S(_{ag})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active: S(<em>{ag}),O(</em>{pt})</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Passive: S(<em>{pt}),Obl(</em>{ag})</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Lummi avoids third person subjects (\(\ast S_3\)):

<table>
<thead>
<tr>
<th>Input: v.ag/3,pt/1</th>
<th>*S(_3)</th>
<th>*S(_{pt})</th>
<th>*S(_{ag})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active: S(<em>{ag}),O(</em>{pt})</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Passive: S(<em>{pt}),Obl(</em>{ag})</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

In languages without passives, the constraint \(\ast S_{pt}\) (or other constraints penalizing passive) are undominated by any of these countervailing constraints.

\(^a\)The analyses of passive here derive from Aissen (1999), but differ in the ranking of constraints (see Bresnan, Dingare, and Manning 2001).
An alternative, historical explanation: actives are basic verb types; passives arise from originally deverbal constructions such as stative adjectives or nominals by a historical process of verbalisation (Trask 1979, Estival and Myhill 1988, Haspelmath 1990, Garrett 1990).

Not sufficient—It begs this question:

Why should actives be basic (underived) verb types synchronically, and not passives?

The intuition: agents make better subjects than patients do. Semantically ‘active’ (proto-agent) arguments tend to receive more attention and are consequently more frequently the subjects of predication. Conversely, semantically ‘inactive’ (proto-patient) arguments tend to receive less individuated attention and are therefore less likely to be subjects of predication.

These tendencies are represented in OT by means of harmonically aligned constraint hierarchies.
A central hypothesis

The same constraints are hypothesized for all grammars, but are more or less active depending on their ranking relative to other constraints.

Lummi falls back on \( *S_{\text{non-topical}} \) with third person agent and patient:

<table>
<thead>
<tr>
<th>input: ( v(\text{ag/3/new,pt/3}) )</th>
<th>( *S_3 )</th>
<th>( *S_{\text{non-topical}} )</th>
<th>( *S_{pt} )</th>
<th>( *S_{ag} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>active: ( S_{ag,Opt} )</td>
<td>*</td>
<td>*!</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>passive: ( S_{pt,Oblag} )</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

In English the person-avoidance constraints are dominated by discourse constraints:

<table>
<thead>
<tr>
<th>input: ( v(\text{ag/3, pt/1}) )</th>
<th>( *S_{\text{non-topical}} )</th>
<th>( *S_{pt} )</th>
<th>( *S_{ag} )</th>
<th>( *S_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>active: ( S_{ag,Opt} )</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>passive: ( S_{pt,Oblag} )</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

We know this because the disharmonic combinations are still grammatical in English, unlike Lummi and Picurí: She met me, She’ll be met by you.
OT provides a model of how the language-particular manifestations of the hierarchies can vary from language to language and even within the same language, while sharing universal trends in direction.
Logical entailment of implicational universals

The theory of harmonic alignment logically entails certain crosslinguistic generalizations, which follow from the central hypothesis of universal constraints, the constraint subhierarchy structure, and the transitivity of constraint domination (\(\Rightarrow\)) in OT. Disregarding other constraints, if passivization is categorical for some input, then it must be categorical for any more marked input.\(^a\)

For example, in Lummi and Picurís, passive is obligatory for input from the lower left cell and optional for input from the lower right cell. Prediction: In no language does the reverse hold.\(^b\)

<table>
<thead>
<tr>
<th>Agent ↓ Patient →</th>
<th>Local person</th>
<th>Third person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local person</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Third person</td>
<td>oblig</td>
<td>opt’l</td>
</tr>
</tbody>
</table>


\(^b\)For discussion of apparent exceptions, see Aissen (2002).
Are the predictions true? —Yes:

Some languages with person/voice interactions:

- **Salishan (Western Canada and USA):**
  - Lushootseed (Coast Salish, Jelinek and Demers 1983)
  - Lummi (Coast Salish, Jelinek and Demers 1983; cf. Klaiman 1991)
  - Squamish (Coast Salish, Jacobs 1994)
  - Bella Coola (isolate, Forrest 1994)

- **Southern Wakashan (Western Canada, USA):**
  - Nootka (Klokeid 1969, Emanatian 1988)
  - Makah (Jacobsen 1979: 156, 159)
  - Nitinat (Klokeid 1978)

- **Kiowa Tanoan (USA):**
  - Picurís (Tiwa-Tewa, Northern Tiwa, Zaharlick 1982)
  - Southern Tiwa (Tiwa-Tewa, Allen and Frantz 1978, Rosen 1990)
  - Arizona Tiwa (Tiwa-Tewa, Tewa, Kroskrity 1985)

- **Mayan (Mexico and Guatemala):** K’iche’ (Mondlach 1981, cited in Aissen 1999)

- **Austronesian: Chamorro (Western Malayo-Polynesian, Guam and Northern Mariana Islands, Chung 1998, Cooreman 1987)**
And no: e.g. Shuswap (Interior Salish, Northern, Kuipers 1974: 47).

<table>
<thead>
<tr>
<th>Object</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 sing.</td>
<td>-cm-, -cl/-cém-, -cél-</td>
</tr>
<tr>
<td>1 plur. incl.</td>
<td>-l/-élm-</td>
</tr>
<tr>
<td>2 sing.</td>
<td>-c/-cí-</td>
</tr>
<tr>
<td>2 plur.</td>
<td>-lm/-úlm-</td>
</tr>
<tr>
<td>3</td>
<td>zero</td>
</tr>
</tbody>
</table>

- s/-és

“The transitive paradigm comprises an (active) Indicative and Imperative, and a Passive (indicative). In the Indicative, passive forms are used for all cases with 1 pl. subject, i.e. we see him, we see you are expressed he is seen, you are seen, etc. . . . ”

Here we have a defective paradigm in the plural category (not uncommon in marked categories—Greenberg 1966). Additional constraints on the morphological expression of number (and other categories) are needed independently. We assume these may in some languages override the constraints on harmonic alignment of person and syntactic function.
Why should we believe that the same constraints are present in every grammar?

The person hierarchy is rooted in cognitive and communicative tendencies which affect not just the formal properties of a few particular languages, but every language.

Is it necessary to assume the constraints are innate?

No, universality does not imply innateness.

Some constraints may reflect innate biases, some may reflect common circumstances of the pragmatic environment.
Recall the Person Hierarchy

1st, 2nd $\succ$ 3rd (local outranks nonlocal)

The Person Hierarchy

— appears at the top of a hierarchy of nominal features: e.g. ‘animacy’, ‘topicality’ hierarchies:
  1st, 2nd $\succ$ 3rd pronominal $\succ$ name $\succ$ human noun $\succ$ animate nonhuman noun $\succ$ inanimate noun

— ranks nominals according to their referents’ “likelihood of participation in the speech event” (Smith-Stark 1974), their “inherent lexical content” (Silverstein 1976), their discourse-pragmatic topicality (Givón 1976, 1979, 1994), or their referents’ accessibility during the psycholinguistic processing of language (Ariel 1990, Warren and Gibson 2001, cf. Gordon, Hendrick, and Johnson 2001)

Note: languages differ in whether first or second person dominates third (DeLancey 1981), but agree on the dominance of first and second over third person.
Two theories of how the person hierarchy influences voice

**perspective-based:** empathy or perspective-taking (Kuno and Kaburaki 1977; DeLancey 1981; Kuno 1987; MacWhinney in progress, ao) — grammar is designed to facilitate perspective shifting during communication; interlocutors share the perspectives of speech-act participants and of referents having causal roles.

**pragmatics-based:** accessibility of referents in the pragmatic context (Givón 1976, 1979, 1994; Ariel 1991; Warren and Gibson 2001; cf. Gordon et al. 2001) — nominal expressions are most easily processed when their referents are contextually accessible

*The connection to voice:* Speech-act participants, referents having causal roles, and contextually accessible referents all tend to receive more attention and are consequently more frequently the subjects of predication.
Can we detect empirically the presence of subordinated person-alignment constraints in the grammar of English?
3. The Stochastic OT Model

Stochastic OT (Boersma 1998)\(^a\) differs from standard OT in two essential ways:

(i) **ranking on a continuous scale:** Constraints are not simply ranked on a discrete ordinal scale; rather, they have a value on the continuous scale of real numbers. Thus constraints not only dominate other constraints, but are specific distances apart, and these distances are relevant to what the theory predicts.

(ii) **stochastic evaluation:** At each evaluation the real value of each constraint is perturbed by temporarily adding to its ranking value a random value drawn from a normal distribution. For example, a constraint with the mean rank of 99 could be evaluated at 98.12 or 100.3. It is the constraint ranking that results from these new values that is used in evaluation.

An OT grammar with stochastic evaluation can generate both categorical and variable outputs.

Categorical outputs arise when crucially ranked constraints are distant. As the distance between constraints increases, interactions become vanishingly rare.\(^b\)

Variable outputs arise when crucially ranked constraints are closer together.

\(^a\)This and subsequent stochastic OT diagrams were created with Praat (Boersma and Weenink 2000).

\(^b\)Units of measurement are arbitrary. With standard deviation = 2.0, a ranking distance of 10 units between constraints is taken to be effectively categorical (Boersma and Hayes 2001: 50).
Generalization to predictions of relative frequency

Disregarding other constraints, if passivization occurs with some frequency for a given input, then (by Aissen’s theory of harmonic alignment expressed within the Stochastic OT model) it must occur with equal or higher frequency for any more marked input (Dingare 2001: 18).^a

<table>
<thead>
<tr>
<th>Agent ↓</th>
<th>Patient →</th>
<th>Local person</th>
<th>Third person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local person</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third person</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

4. Statistical Person/Voice Interactions in English

To see if there is an effect of person on the selection of active or passive in English, we need information about the systemic choices made for each input. Prior studies generally fail to provide the full joint distribution, from which we can reconstruct the conditional frequencies needed.\(^a\) We have therefore examined the parsed SWITCHBOARD corpus, a database of spontaneous telephone conversations spoken by over 500 American English speakers, both male and female, from a great variety of speech communities (Godfrey et al. 1992). The conversations average 6 minutes in length, collectively amounting to 3 million words. We have used the parsed portion of this corpus (released as part of the Penn Treebank, Marcus et al. 1993), which contains 1 million words.

Although the frequency of passives is quite low in this corpus, the frequency of local pronouns is high.\(^b\)

\(^a\)Estival and Myhill (1988) provide exactly the kind of information needed for animacy and definiteness, but they provide person frequencies only for the patient role. Weiner and Labov (1981) study the frequency of choice between an agentless passive (We’re not allowed to smoke) and an equivalent generalized-subject active (They don’t allow us to smoke”), but do not study full passives.

\(^b\)Francis, Gregory, and Michaelis (1999) show that 91% of subjects in the parsed SWITCHBOARD corpus are pronominal.
Proportion of Full Passives to Passivizable Transitives in SWITCHBOARD

\[
\frac{p}{(p+a)}
\]

<table>
<thead>
<tr>
<th>Agent ↓</th>
<th>Patient →</th>
<th>Local person</th>
<th>Third person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local person</td>
<td>0.0%</td>
<td>0.0%</td>
<td></td>
</tr>
<tr>
<td>Third person</td>
<td>2.9%</td>
<td>1.2%</td>
<td></td>
</tr>
</tbody>
</table>

Compared to the rate of passivization for inputs of third persons acting on third persons (1.2%), the rate of passivization for first or second person acting on third is substantially depressed (0%) while that for third acting on first or second (2.9%) is substantially elevated.

Harmonic alignment gave us two particular hypotheses which are supported by these data: that the rate of passivization of 3 → 1,2 should be higher than for 3 → 3 (1-sided Fisher exact, \( p < 0.008 \)); and that the rate of passivization of 1,2 → 3 should be lower than for 3 → 3 (1-sided Fisher exact, \( p < 0.0001 \)).
The stochastic generalization:

The same disharmonic person/argument associations which are avoided categorically in languages like Lummi and Picurís by making passives either impossible or obligatory, are avoided in the SWITCHBOARD corpus of spoken English by either depressing or elevating the frequency of passives relative to actives.

---

5. Stochastic Grammars

Partial stochastic grammar of English:

Partial stochastic grammar of Lummi:
Observe: \( *S_{pt} \gg *S_3 \) but \( |*S_{pt} - *S_3| = 6 \), close enough to produce low frequency variable outputs for some inputs. For inputs where only the agent is third person, passive outputs will occasionally be favored by \( *S_3 \):

An (infrequent) effect of \( *S_3 \) on passive outputs:

<table>
<thead>
<tr>
<th>input: ( v(\text{ag}/3,\text{pt}/1) )</th>
<th>( *S_3 )</th>
<th>( *S_{pt} )</th>
<th>( *S_{ag} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>active: ( S_{ag},O_{pt} )</td>
<td>!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>passive: ( S_{pt},\text{Obl}_{ag} )</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When both agent and patient are third person, the \( *S_3 \) constraint cannot decide between active and passive, and the decision passes to other constraints.
Observe: |*Obl\textsubscript{1,2} − *O\textsubscript{pers}| > 10. (*O\textsubscript{1,2} disfavors an active for an input with local-person patient and *O\textsubscript{3} for an input with third-person patient.) These rankings reflect the zero frequency of local person passive agents in the data. See Kuno and Kaburaki (1977), Kuno (1987: 230–231): *And then I hit John on the head*, *And then John was hit on the head by me*. But Kato (1979) cites (from Studs Terkel, *Working*):

I said, “Me watch it! Fuck that! Let him watch it.” He was hired by me. I could fire him if I didn’t like him.

When somebody says to me, “You’re great, how come you’re just a waitress?” *Just* a waitress. I’d say, “Why, don’t you think you deserve to be served by me?”
Caveats:

- With more training data and a more complete constraint set which includes factors of topicality and focus, the model should learn grammars that produce passives with local person agents.\(^a\)

- The set of constraints used in this system is motivated by broader typological considerations. Some of these constraints play no necessary part in the system presented here, and a smaller constraint set is able to model the observed data equally well.

- This constraint set contains no information structure constraints which would motivate the use of passive independent of person. Because of this, the grammar models the ‘background level’ of passivization by keeping \(*S_{ag}\) close enough to \(*S_{pt}\) that one will occasionally get passives. This can be viewed as an artifact of the incomplete constraint set.

\(^a\)If the ranking value of \(*\text{Obl}_{1,2}\) in the grammar were lowered from 109 to 104, the output of local person passives would increase to one-tenth of one percent, 0.1\%, while barely changing the frequency of other outputs.
Lummi...

Unfortunately we lack a parsed SWITCHBOARD corpus for Lummi or Picurís. Nevertheless, it is possible to show by simulation how the descriptions of passive/voice interactions in Lummi or Picurís grammar can also be captured by a stochastic OT grammar. We interpret the descriptions of Lummi from Jelinek and Demers (1983, 1994) by means of a simple distribution.

*Where a sentence type is described as ungrammatical, we assign it 0% relative frequency; where it is described as obligatory, we assign it 100%; and where it is described as optional, we assign it 50%.*

<table>
<thead>
<tr>
<th>Lummi (Jelinek and Demers 1983, 1994):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascribed rate of Passivization</td>
</tr>
<tr>
<td>Agent ↓ Patient → Local person Third person</td>
</tr>
<tr>
<td>Local person 0.0% 0.0%</td>
</tr>
<tr>
<td>Third person 100% 50%</td>
</tr>
</tbody>
</table>
Observe: \(|*_{S_3} - *_{S_{pt}}| > 10\). This ranking yields the obligatory passivization of inputs with local person patients and non-local person agents, capturing the categorical influence of person on Lummi passivization.\(^a\)

\(^a\)This analysis, deriving from GLA simulations of Bresnan, Dingare, and Manning (2001), differs from that of Aissen (1999), though the constraints are the same.
Conventionalization and Frequency

Stochastic OT grammars allow us to place the person/voice interactions in English and Lummi at points on a continuum of conventionalization that connects frequentistic preferences in usage to categorical grammatical constraints. If this general perspective is correct, then we would expect to find languages at intermediate points on this same continuum.

Consider Squamish:
3 → 2: passive obligatory in Lummi and Squamish
3 → 1: passive obligatory in Lummi, optional in Squamish

Analysis:

Lummi:                Squamish:
*Obl₁,₂ ≫ *S₃ ≫ *O₂,*O₁,*Sₚᵗ    *Obl₁,₂ ≫ *O₂ ≫ *S₃,*O₁,*Sₚᵗ

---

"Recall that the mutual ranking of the local-person avoidance constraints is not fixed by the subhierarchy, but subject to crosslinguistic variation."
However, it is not fully informative to say, as has been customary (Jelinek and Demers 1983 ao), that passivization with third person agents and first person patients is “optional” in Squamish.

In terms of what is preferred rather than what is merely possible, Squamish is described as being much the same as Lummi, “except that third person acting on first may be active, though commonly passive” (Klokeid 1969: 11).\(^a\)

Thus in Squamish as in English, passives of the type *I was fooled by her* are optional alternatives to actives with disharmonic local-person objects: *She fooled me*. But in spoken English, such passives are exceedingly infrequent, far less common than the corresponding actives, while in Squamish they appear to be more frequent than the corresponding actives. Why?

\(^a\)Bresnan, Dingare, and Manning were unable to find quantitative measures of Squamish passives by person. Jacobs’ (1994) corpus study of Squamish excludes first and second person because the purpose is to examine interactions of topic continuity with voice/inversion through measures of distance between pronouns and their textual antecedents.
Squamish and Lummi are closely related Coast Salish languages. In the continuous constraint space of stochastic OT, the similarities of their grammars to each other and to the grammars of their common ancestors will appear as close distances between constraints.

In particular, different points in the changing categoricity of person effects on the passive will be reflected by gradual changes in frequency, as the relative distance between constraints shrinks and grows:

Smooth Lummi-Squamish Reranking:

\[ *_{Ob1,2} \gg *_{S_3} \gg *_{O_2, O_1, S_{pt}} \]

---

\(^a\)The rankings of Aissen (1999) differ somewhat from those learned by the GLA, though the sets of possible outputs are equivalent.
Reranking produces smooth changes in frequency—

If reranking is the movement in strength of a constraint along the continuous scale, as implied by the stochastic OT model, then (all else being equal) smooth changes in the relative frequencies of usage are predicted.

—but not linear changes:

If a constraint reranking is crucial to the choice between two outputs, and the distance between the two constraints is changing linearly, the prediction is that we should see an ‘S’ curve between the proportion of occurrences of the two outputs, of the sort that has been widely remarked on in historical and socio-linguistics (Weinreich, Labov, and Herzog 1968, Bailey 1973, Kroch 2001).
Logistic response

Proportion of the time output is passive vs. Difference in base constraint ranking

English

Squamish

Lummi
“Not all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity.”

— Weinreich, Labov, and Herzog (1968: 188)
6. Constraint Overlap and Statistical Dependencies of the Input

Because of statistical dependencies of the input, other constraints can also derive a soft person effect.
Previous studies show that givenness/topicality interacts with the choice of active or passive. Among others:

*Estival and Myhill (1988: 457–8)*:

The relative frequency of the choice of passive over active increases with ‘nontopical’ (nominal, nonhuman, and indefinite) agents and with ‘topical’ (pronominal, human, or definite) patients.

*Weiner and Labov (1981: 46–54)*:

The relative frequency of the choice of passive over active increases with given patients and decreases with new patients. ['Given’ arguments are operationally defined to have a coreferential NP within the preceding 5 clauses, irrespective of speaker turns.]
Hypothetical StOT grammar for the English-style ‘pragmatic passive’:\(^a\)

\(\ast S_{pt}\) \(\ast S_{\text{newer}}\) \(\ast S_{t}\) \(\ast O_{1,2}\)

<table>
<thead>
<tr>
<th>Strict</th>
<th>104</th>
<th>99.6</th>
<th>90.1</th>
<th>Lax</th>
</tr>
</thead>
</table>

\[\begin{array}{|c|c|c|c|}
\hline
\text{input: } v(\text{ag/new, pt}) & \ast S_{pt} & \ast S_{\text{newer}} (\text{or } \ast S_{t}) & \ast O_{1,2} \\
\hline
\text{active: } S_{ag, O_{pt}} & & & \\
\text{passive: } S_{pt,Obl_{ag}} & \ast! & & \\
\hline
\end{array}\]

\[\begin{array}{|c|c|c|c|}
\hline
\text{input: } v(\text{ag/new, pt}) & \ast S_{\text{newer}} (\text{or } \ast S_{t}) & \ast S_{pt} & \ast O_{1,2} \\
\hline
\text{active: } S_{ag, O_{pt}} & \ast! & & \\
\text{passive: } S_{pt,Obl_{ag}} & & \ast & \\
\hline
\end{array}\]

\[^a\ast S_{\text{newer}} = \text{avoid subjects which are discourse-newer than non-subject arguments of the same clause (Birner and Ward 1998) (} \approx \text{ Aissen’s (1999) } \ast \text{Su/x).}\]
How topicality/accessibility derives a soft person effect.

Given that first and second persons are seldom discourse new, while third persons may be (Cooreman 1987), one could assume that local person subjects are not penalized by the avoidance of newer subjects, while non-local person subjects are. Then a soft effect of the person hierarchy would follow from $S_{newer}$ in addition to any effect of person-avoidance constraints.

Nonlocal agents are differentially favored for passivization by newness (assuming that third person is inherently newer than local person):

<table>
<thead>
<tr>
<th>Input: $v(\text{ag}/3/\text{new}, \text{pt}/3)$</th>
<th>$S_{newer}$ (or $S_t$)</th>
<th>$S_{pt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>active: $S_{ag}, O_{pt}$</td>
<td>$!$</td>
<td></td>
</tr>
<tr>
<td>passive: $S_{pt}, Obl_{ag}$</td>
<td></td>
<td>$*$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input: $v(\text{ag}/2, \text{pt}/3)$</th>
<th>$S_{newer}$ (or $S_t$)</th>
<th>$S_{pt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>active: $S_{ag}, O_{pt}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>passive: $S_{pt}, Obl_{ag}$</td>
<td>$!$</td>
<td>$*$</td>
</tr>
</tbody>
</table>
Similarly, nonlocal patients are differentially disfavored for passivization by newness (again assuming that third person is inherently newer than local person):

<table>
<thead>
<tr>
<th>Input: (v(\text{ag/3, pt/3/new}))</th>
<th>(*S_{\text{newer}}) (or (*S_t))</th>
<th>(*S_{\text{pt}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active: (S_{\text{ag}}, O_{\text{pt}})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passive: (S_{\text{pt}}, O_{\text{blag}})</td>
<td>(*!)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input: (v(\text{ag/3, pt/2}))</th>
<th>(*S_{\text{newer}}) (or (*S_t))</th>
<th>(*S_{\text{pt}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active: (S_{\text{ag}}, O_{\text{pt}})</td>
<td>(*!)</td>
<td></td>
</tr>
<tr>
<td>Passive: (S_{\text{pt}}, O_{\text{blag}})</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Soft ‘Topicality’ Effects on English Active/Passive Choice: the Switchboard Corpus

The parsed Switchboard corpus is not tagged for topicality or givenness, but we can approximate this information-status concept by comparing the distributions of more and less definite nominal expression types, such as pronouns, proper names, definite and indefinite noun phrases.

In the Treebank Switchboard corpus local person pronouns are plentiful, but the distribution of pronouns and lexical (= nonpronominal) NPs is highly skewed (Francis et al. 2001):

- 91% of subjects are pronominal
- 66% of objects are lexical

There are significant ‘topicality’ effects on passivization in Switchboard, following the methods of Dingare (2001: 19–23). \(^{b}\)

---


\(^{b}\)The results reported here are unpublished work in progress by Bresnan.
Switchboard Rates of Passive by Pronominality of Agent and Patient\textsuperscript{a}

<table>
<thead>
<tr>
<th>Pt → Pronoun</th>
<th>Non-Pronoun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag ↓ Pronoun</td>
<td>A: 2457</td>
</tr>
<tr>
<td></td>
<td>P: 4 (0.16%)</td>
</tr>
<tr>
<td>Non-Pronoun</td>
<td>A: 179</td>
</tr>
<tr>
<td></td>
<td>P: 18 (9.14%)</td>
</tr>
</tbody>
</table>

A: Active count; P: Passive count; \((n\%):\) percent passives

\[\text{: one-sided Fisher exact, } p < 0.05\]

\[\text{: one-sided Fisher exact, } p < 0.0001\]

\textsuperscript{a}Pronouns = definite personal pronouns and reflexives; Non-Pronouns = Definites, Proper Names, and Indefinites.
Switchboard Rates of Passive by Definiteness of Agent and Patient\(^a\)

<table>
<thead>
<tr>
<th>Pt → All Definite</th>
<th>Indefinite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag ↓</td>
<td></td>
</tr>
<tr>
<td>All Definite</td>
<td>A: 4756</td>
</tr>
<tr>
<td></td>
<td>P: 23 (0.48%)</td>
</tr>
<tr>
<td>Indefinite</td>
<td>A: 42</td>
</tr>
<tr>
<td></td>
<td>P: 7 (14.29%)</td>
</tr>
</tbody>
</table>

A: Active count; P: Passive count; \((n\%): \text{percent passives}

: one-sided Fisher exact, \(p < 0.05\)

: one-sided Fisher exact, \(p < 0.0001\)

\(^a\)All Definites = Pronouns, Proper Names, and Definites. Definites = NPs beginning with *the*, *this*, *that*, *these*, or *those* and not followed by a proper noun; Indefinites = NPs beginning with *a*, *an*, or *some* and not followed by a proper noun. NPs beginning with possessives are excluded.
Hard ‘Topicality’ Effects in Lummi

Lummi categorically avoids pronominal objects with non-pronominal subjects (Jelinek and Demers 1983, 1994):

* __ ‘The man knows it.’

\[
\text{\text竝 ki-t-}\quad \circ\quad \text{c}\quad \text{swey?qe?} \\
\text{know-TR-PASS by the man}
\]

‘It is known by the man’

\[
\text{\text竝 ki-t-s}\quad \circ\quad \text{c}\quad \text{swey?qe?} \\
\text{know-TR-3.TR.SUBJ the man}
\]

‘He knows the man.’

These and related facts\textsuperscript{a} follow from the theory of harmonic alignment in OT: constraints favoring the harmonic association of referentially prominent arguments (pronoun, definite) with prominent syntactic function (subject) are hypothesized to be present as subhierarchies of the grammars of all languages, and to yield categorical grammaticality effects if they outrank other relevant constraints (Aissen 1999).

The stochastic generalization of OT explains how these categorical topicality effects of Lummi grammar can parallel the soft, frequentistic effects of ‘topicality’ on voice in English, which lie beneath the threshold of grammaticality judgments.

\textsuperscript{a}The avoidance of local person objects and passive agents does not extend to the free-standing referring expressions for local persons, which are focussed, hence ‘newer’ (Jelinek and Demers 1994: 714).
Recall Givón once again:

“What we are dealing with is apparently the very same communicative tendency—to reserve the subject position in the sentence for the topic, the old-information argument, the “continuity marker.” In some languages (Krio, etc.), this communicative tendency is expressed at the categorial level of 100%. In other languages (English, etc.) the very same communicative tendency is expressed “only” at the noncategorial level of 90%. And a transformational–generative linguist will then be forced to count this fact as competence in Krio and performance in English.”

— Givón (1979: 26–31)
Does that mean that person constraints are not needed to explain the interaction of person with voice?

No. Recall Squamish:

3 → 2: passive obligatory in Lummi and Squamish, optional in English
3 → 1: passive obligatory in Lummi, optional in Squamish and English

There is no independent reason to believe that the Speaker is systematically less ‘given’ than the Hearer in Squamish vs. Lummi, or that the Hearer is systematically more ‘given’ than the Speaker in Squamish vs. English.

Conclusion: person-avoidance is controlled independently of information structure.

Nevertheless, we expect them to overlap substantially . . . (Dingare 2001)
Statistical dependencies of the input cause overlapping constraints to rise and fall together under the Gradual Learning Algorithm.

The fact that two constraints have exactly the same violation marks on a given input means that during training, they will be treated the same (that is, demoted or promoted by the same amount). If the given input is frequent, a person constraint will end up close enough to an overlapping discourse constraint to drive the choice of the output to some degree even when the candidates are not driven by topicality. Hence, because a local person agent is rarely realized as an oblique, the speaker may disprefer passive even when the local person agent is non-topical.

Thus, because of the statistical dependencies of person and ‘topicality’ in the input, the person and discourse constraints will rise in tandem under the GLA. In the absence of active countervailing constraints, an emerging categoricity of person effects on voice will necessarily accompany an emerging categoricity of newness effects on voice (*$S_{newer}$), and vice versa.
7. Questions about the passive analysis

Question: Where is the rest of the grammar?

Answer: There are further lexical, morphosyntactic, and syntactico-semantic optimizations, for which we must choose a specific representational basis.

One such representational basis is OT-LFG, which is formally well defined and computationally implemented.\(^a\)

---

/uts.cc.utexas.edu/~jonask/ for related work.
Parallel Optimizations (OT-LFG):

[Diagram of syntactic tree with annotations for each node, showing the structure of sentences with various grammatical relations and lexical items.]
**Question:** How can actives and passives be competing expressions of the same input when the constructions may differ in meaning?

(a) Everyone likes someone.
(b) Someone is liked by everyone.  

(c) Reluctantly, Joan instructed Mary.
(d) Reluctantly, Mary was instructed by Joan.  

(Chomsky 1957: 100–1)

(c) Reluctantly, Joan instructed Mary.
(d) Reluctantly, Mary was instructed by Joan.  

(McConnell-Ginet 1982)

**Answer:** Subject quantifiers preferentially scope over nonsubject quantifiers. Since subject selection differs in the active and passive, the preferred interpretations with quantifiers also differs in (a)–(b).

Some manner adverbs modify either subjects or agents. Since subjects and agents coincide in actives and diverge in passives, the possible adverb interpretations also differ in (c)–(d).

In short, actives and passives differ in argument realization in the syntax, and these differences interact with the interpretation of quantifiers and adverbs.
The interactions of semantics with subject realization would be built into a more extensive optimization system. For example, if the input includes a specific semantic contrast in quantifier scope, the speaker could optimize the syntactic choice of active or passive to preserve the preferred correspondence between quantifier scope and linear precedence (or syntactic prominence), unless other constraints overrode it.

In simple sentences lacking such quantifiers and adverbs, actives and passives are often semantically equivalent, so that other factors such as information structure and person may occasionally determine the choice.

In Lummi and Picurís, passives fill gaps in the active paradigm created by antiharmonic person combinations, and vice versa.

In Tzotzil and Chamorro, passives fill gaps in the active paradigm created by antiharmonic animacy combinations, and vice versa (Aissen 1999).

Conclusion: Use and meaning can be elaborated and differentiated by further constraints, but the sets of inputs for actives and passives in simple sentences overlap in English and some other languages.
Could competing conventional generative grammars explain the passive variation in English?

The competing grammars theory of variation is a model of diglossia (Kroch 2001). On the diglossic model of variation, the contact between two different populations having different grammars leads to internalization of competing grammars by individual speakers, who control two separate varieties. For example, some historical changes in English word order are attributed to the influence of Scandinavian speakers in Northern England (Kroch and Taylor 1997).

Could the diglossic model explain the passive findings? On this account, individual speakers would vary in the frequency of passive outputs because they have internalized alternative grammars which they deploy with varying frequency. The different grammars would have arisen from contact between different populations speaking varieties of English with and without the passive construction for certain person/role combinations. One population would have Lummi-like gaps in actives and passives as a hard constraint of their English grammar, perhaps as a result of some parameter setting of UG.
Some early studies propose that middle-class English speakers use an ‘elaborated code’ which has a higher proportion of passive verbs among all finite verbs than a ‘restricted code’ of working-class speakers, which has a lower percentage (Bernstein 1971 ao). But these studies have been criticized for failing to isolate the syntactic choice between active and passive, which shows no significant difference between these groups (Weiner and Labov 1981: 32). *(Passives should be compared to equivalent actives, rather than to all sentences. The latter can be influenced by differences in what is talked about, given that passives require fewer arguments than actives.)*

Spontaneous speech shows significant stylistic and discourse effects on the choice of (agentless) passive or generalized-subject active. But:

“All of these conditions on the selection of active vs. passive are general features of the English language, used in much the same way by the very different sub-section of the speech communities that we studied.” *(Weiner and Labov 1981: 56)*

**Conclusion:** Diglossia is an unlikely model for the passive data.

---

*a*Generalized pronoun subjects (“they”) are characteristic of colloquial English, while passives are a mark of formal scientific and literary discourse; passives are also favored by the discourse tendencies to preserve subject reference and structural parallelism.