1

An axiomatic approach to speaker preferences

In this paper, we propose a set of axioms that can be used to model the choices made by speakers and hearers when pairing speech signals with meanings in speech production and understanding. We then prove that if these axioms are a valid characterization of speakers and hearers, then speaker and hearer preferences can always be encoded by means of a measurable utility. Finally, we discuss the potential ramifications for linguistic theory.

1.1 Introduction

Both verbal and non-verbal human communication can be viewed as a sequence of exchanges between speakers and hearers via sound, text, or gesture. In each exchange, a speaker uses speaker language L and speaker context C to encode intended meaning M as speech signal S. The speech signal S, after being distorted by noise, is then perceived as speech signal S' by a hearer, who tries to guess the speaker's intended meaning by using hearer language L' and hearer context C' to construct a perceived meaning M'. The exchange is shown schematically below:

$$M \xrightarrow[L,C]{\text{production}} S \xrightarrow{\text{noise}} S' \xrightarrow[L',C']{\text{understanding}} M'$$

We will refer to the tuples $A = (S, M, C)$ and $A' = (S', M', C')$ as the *analyses* made by the speaker and hearer during the communication.

Communication involves a substantial amount of guessing and uncertainty. Speakers must try to find the analysis $A = (S, M, C)$ that

FG-2006:
The 11th conference on Formal Grammar.
TBA.

they think best conveys the meaning M to the hearers, given speaker language L and context C (the problem of *production* or *generation*). Similarly, hearers must try to find the analysis $A' = (S', M', C')$ that they think fits best with the noisy speech signal S' originating from the speaker, given hearer language L' and context C' (the problem of *understanding* or *parsing*). In order to find the analysis that seems best in terms of facilitating the communication, hearers and speakers must obviously be capable of comparing different analyses with each other, ie, their preferences must define an ordering on the space of all conceivable analyses, which includes the ungrammatical analyses that hearers construct in the presence of speaker errors. From a linguistic perspective, it is therefore essential to ask about the nature of the ordering induced by these preferences.

Different linguistic theories provide different answers to this question. For example, in probabilistic language models, the ordering is defined in terms of a probability measure, and in OT, the ordering is defined by counting the violations of differently ranked OT constraints. However, rather than exploring some of the different ways in which the ordering can be defined, we will take a more principled approach. First, we will propose an axiomatic model of speaker preferences based on axioms that we believe must be satisfied by all linguistically reasonable models of speaker preferences. We will then prove that speaker preferences that satisfy these axioms can be encoded by means of a measurable utility, ie, there always exists an order-preserving mapping from the space of all analyses to the set of real numbers for such speaker preferences. Finally, we will discuss the potential ramifications for linguistic theory.

1.2 An axiomatic model of speaker preferences

When modelling speaker preferences, we believe that it is a reasonable idealization to assume that speakers are not only capable of comparing isolated analyses, but that they are also capable of comparing arbitrary unordered collections of isolated analyses, which we will refer to as *corpora*. It is important to note that in our definition of corpora, the individual analyses in a corpus are unrelated and unordered, but individual analyses may well be analyses of entire texts, so that the model is capable of capturing structural dependencies between different parts of the same text.

Language teachers routinely compare their students on the basis of their linguistic performance in a series of unrelated essays, and as speakers, we seem to have intuitions about some speakers being better at expressing themselves than others. That is, as speakers, we gladly make the comparisons that are the basis of our idealization, so the idealization is not far removed from real-world practice.

In our axiomatic model of speaker preferences, we assume that the speaker's ordering of the corpora satisfies the axioms C0–C4 below (for all corpora a, b, c, d and positive integers k). In the axioms, we use the following notation. Given two corpora a and b, we let $a + b$ denote the concatenation of the two corpora, and we let ka denote the corpus consisting of k copies of a where k is a positive integer. Moreover, we use the notation $a \succ b$ if the agent prefers a to b, $a \sim b$ if the agent is indifferent between a and b, and $a \succsim b$ if the agent prefers a to b or is indifferent between them.

> **Axiom C0**. The speaker's ordering is complete.

That is, given a choice between two different corpora, the speaker will always prefer one over the other, or be indifferent between them, and the resulting ordering is transitive.

> **Axiom C1**. $a \succsim b$ iff $a + c \succsim b + c$.

That is, the speaker's relative ordering of two corpora only depends on the analyses that are not contained in both of them, and we can therefore add or delete a shared subcorpus without affecting the ordering.

> **Axiom C2**. $a \succsim b$ iff $ka \succsim kb$.

That is, if the speaker prefers a over b, then the speaker also prefers k copies of a over k copies of b, and vice versa, ie, the ordering is unaffected by scaling the two corpora up or down.

> **Axiom C3**. If $a \succsim b$ and $c \succsim d$, then $a + c \succsim b + d$.

That is, if the speaker prefers all the parts of a corpus when making a part-wise comparison with another corpus, the speaker also prefers the entire corpus.

> **Axiom C4.** If $a \succ a'$ and b is any corpus, then there exists a positive integer k such that $ka \succ ka' + b$.

That is, if a is better than a', then we can always find k such that the difference in goodness between k copies of a and k copies of a' is larger than the goodness of any given b. Or, phrased differently, a sufficiently large number of small errors will always outweigh a large error.

Is it a reasonable idealization to assume that human speakers are capable of comparing corpora in a way that satisfies Axioms C0–C4? Given the limitations of the human mind, it would be surprising if real speakers were capable of comparing arbitrarily large corpora, or of using Axioms C0–C4 consistently. However, both probabilistic language models and OT naturally lead to an ordering of corpora that satisfies Axioms C0–C3, and we find it hard to imagine a linguistically reasonable ordering that fails to satisfy these axioms. Axiom C4 is satisfied by probabilistic language models, but not by OT, since the violation of a high-ranked constraint always outweighs any number of violations of lower-ranked constraints in OT. However, this is a point where OT has been criticized: the psycholinguistic evidence about "ganging-up effects" seems to suggest that a large error can always be outweighed by a sufficient number of small errors, ie, Axiom C4 seems to hold (cf. Sorace and Keller, 2005, §2.3).

1.3 Speaker preferences as measurable utilities

We will now show that if the reader accepts our axiomatic model of speaker behaviour, then the reader is also forced to accept that a speaker's preference ordering can be expressed by means of a measurable utility. That is, in any linguistic model of human communication that is compatible with our axioms, measurable utilities are a necessity and not just a possibility. In our proof, we will show that we can extend the speaker's ordering of corpora to an ordering on a mixture set that satisfies the axioms of utility. By invoking the von Neumann-Morgenstern theorem, we can then show that the ordering can be expressed by means of a measurable utility.

First, we need to define the notions of complete ordering, mixture set, and measurable utility (cf. Herstein and Milnor (1953); for an informal overview of utility theory, see Russell and Norvig (1995, pp. 473–475)).

Definition 1 A *complete ordering* (or *total ordering*) defined on a set $\mathcal{S}$ is a binary relation $\succsim$ on $\mathcal{S} \times \mathcal{S}$ such that (i) for any $a, b \in \mathcal{S}$, we have $a \succsim b$ or $b \succsim a$, and (ii) for all $a, b, c \in \mathcal{S}$, if $a \succsim b$ and $b \succsim c$, then $a \succsim c$. If $a \succsim b$ and $b \succsim a$, we write $a \sim b$ and say that a and b are *indifferent*. If $a \succsim b$ and a and b are not indifferent, we say that a is *preferred over* b and write $a \succ b$.

Definition 2 Let $\mathcal{S}$ be a set equipped with a function that given any $a, b \in \mathcal{S}$ and $\lambda \in [0, 1]$ returns an element in $\mathcal{S}$, which we denote by $\lambda a + (1 - \lambda)b$ and call a *mixture* of a and b. We say that $\mathcal{S}$ is a *mixture set* if $\mathcal{S}$ satisfies the following three conditions for all $a, b \in \mathcal{S}$ and $\lambda, \mu \in [0, 1]$:

(i) $1a + 0b = a$
(ii) $\lambda a + (1 - \lambda)b = (1 - \lambda)b + \lambda a$
(iii) $\lambda(\mu a + (1 - \mu)b) + (1 - \lambda)b = \lambda\mu a + (1 - \lambda\mu)b$

Definition 3 Let $\mathcal{S}$ be a mixture set with a complete ordering $\succsim$. A function $u\colon \mathcal{S} \to \mathbb{R}$ is called a *measurable utility* if u is order-preserving and linear, ie, u satisfies:

(i) $u(a) > u(b)$ if and only if $a \succ b$, for all $a, b \in \mathcal{S}$;
(ii) $u(\lambda a + (1-\lambda)b) = \lambda u(a) + (1-\lambda)u(b)$ for all $a, b \in \mathcal{S}$ and $\lambda \in [0,1]$.

The von Neumann-Morgenstern theorem was discovered independently by von Neumann and Morgenstern (1944) and Ramsey (1931). Herstein and Milnor (1953) sharpened the result and gave an elegant proof, using a more general set of axioms. The following formulation of the theorem is due to Herstein and Milnor.

Theorem 1 (von Neumann and Morgenstern) *Let $\mathcal{S}$ be a mixture set with a complete ordering $\succsim$. Then a measurable utility can be defined on $\mathcal{S}$ if and only if $\mathcal{S}$ satisfies the following two axioms:*

Axiom 1′ (weak substitutability). *If $a, a' \in \mathcal{S}$ and $a \sim a'$, then $\frac{1}{2}a + \frac{1}{2}b \sim \frac{1}{2}a' + \frac{1}{2}b$ for all $b \in \mathcal{S}$.*
Axiom 2′ (continuity). *The sets $\{\lambda \mid \lambda a + (1-\lambda)b \succsim c\}$ and $\{\lambda \mid c \succsim \lambda a + (1-\lambda)b\}$ are closed for all $a, b \in \mathcal{S}$.*

We will now formalize our notion of a corpus.

Definition 4 Let $\mathcal{A}$ be a set of analyses, and let S be a subset of $\mathbb{R}$ that contains 0 and 1 while being closed under addition and multiplication. A *corpus* over S is a function $c\colon \mathcal{A} \to \mathcal{S}$ such that $c(a) = 0$ for all but finitely many a. The set of all corpora $c\colon \mathcal{A} \to \mathcal{S}$ is called the *corpus set* induced by $\mathcal{A}$ over S.

Our axiomatic model of speaker preferences corresponds to the case where S is the set $\mathbb{N}_0$ of all non-negative integers. However, in our proof, we will extend this initial corpus set by replacing S with the set $\mathbb{Q}$ of rational numbers, and the set $\mathbb{R}$ of real numbers.

Definition 5 We define addition and scalar multiplication on a corpus set $\mathcal{C}$ by letting $c + c'$ denote the corpus that maps $a \in \mathcal{A}$ to $c(a) + c'(a)$, and letting sc denote the corpus that maps $a \in \mathcal{A}$ to $sc(a)$, given $c, c' \in \mathcal{C}$ and $s \in S$. We let 0 denote the corpus that maps all $a \in \mathcal{A}$ to 0, and note that $c + 0 = c$ and $1c = c$ for all $c \in \mathcal{C}$.

We can now restate our axioms for speaker preferences for corpora with arbitrary $\mathcal{S}$.

Definition 6 Let $\mathcal{C}$ be a corpus set with a complete ordering $\succsim$. We say that $\mathcal{C}$ is a *corpus space* if $\mathcal{C}$ satisfies the following four axioms for all $a, b, c, d \in \mathcal{C}$ and positive integers k:

Axiom C1. $a \succsim b$ iff $a + c \succsim b + c$.

Axiom C2. $a \succsim b$ iff $ka \succsim kb$.
Axiom C3. If $a \succsim b$ and $c \succsim d$, then $a + c \succsim b + d$.
Axiom C4. If $a \succ a'$ and $b \in \mathcal{C}$, then there exists a positive integer k such that $ka \succ ka' + b$.

Obviously, our axiomatic model of speaker preferences corresponds to saying that the speaker has an ordering on the corpus set $\mathcal{C}$ induced by $\mathcal{A}$ over $\mathbb{N}_0$ that turns $\mathcal{C}$ into a corpus space. We will now show that $\mathcal{C}$ can be extended to corpus spaces $\mathcal{C}_\mathbb{Q}$ and $\mathcal{C}_\mathbb{R}$ over $\mathbb{Q}$ and $\mathbb{R}$, respectively.

Definition 7 Let $\mathcal{C}_\mathbb{Q}$ be the corpus set induced by $\mathcal{A}$ over $\mathbb{Q}$. Given $c \in \mathcal{C}_\mathbb{Q}$, let p_c denote the positive part of the corpus c, ie, the corpus that maps $a \in \mathcal{A}$ to $\max(0, c(a))$, and let q_c denote the smallest positive integer such that $q_c c(a)$ is an integer for all $a \in \mathcal{A}$. Define an ordering $\succsim$ on $\mathcal{C}_\mathbb{Q}$ by

$$a \succsim b \qquad \text{iff} \qquad q_a q_b (p_a + p_{-b}) \succsim' q_a q_b (p_{-a} + p_b)$$

where $\succsim'$ denotes the ordering on $\mathcal{C}$.

Proposition 2 *$\mathcal{C}_\mathbb{Q}$ is a corpus space over $\mathbb{Q}$ that contains $\mathcal{C}$, and the ordering on $\mathcal{C}_\mathbb{Q}$ coincides with the ordering on $\mathcal{C}$.*

Proof. Since $p_c = c$, $p_{-c} = 0$, and $q_c = 1$ for all $c \in \mathcal{C}$, it follows immediately that $\succsim$ coincides with $\succsim'$ on $\mathcal{C}$. To show that $\succsim$ is well-ordered, it suffices to show that $a \succsim b$ or $a \precsim b$ for all $a, b \in \mathcal{C}_\mathbb{Q}$. So suppose $a \succ b$ and $a \prec b$, then $q_a q_b (p_a + p_{-b}) \succ' q_a q_b (p_{-a} + p_b)$ and $q_a q_b (p_a + p_{-b}) \prec' q_a q_b (p_{-a} + p_b)$, a contradiction. To show that $\succsim$ is transitive, note that $a \succsim b$ and $b \succsim c$ implies $q_a q_b (p_a + p_{-b}) \succsim' q_a q_b (p_{-a} + p_b)$ and $q_b q_c (p_b + p_{-c}) \succsim' q_b q_c (p_{-b} + p_c)$. Thus, by Axiom C1 we can add $q_a q_b p_{-c}$ to the first inequality and $q_b q_c p_{-a}$ to the second inequality, use Axiom C2 to multiply the first inequality with q_c and the second with q_a, and then use transitivity in $\mathcal{C}$ to conclude:

$$\begin{aligned} q_a q_b q_c (p_a + p_{-b} + p_{-c}) &\succsim' q_a q_b q_c (p_{-a} + p_b + p_{-c}) \\ &\succsim' q_a q_b q_c (p_{-a} + p_{-b} + p_c). \end{aligned}$$

By Axiom C1, we can now subtract $q_a q_b q_c p_{-b}$, giving $q_a q_b q_c (p_a + p_{-c}) \succsim' q_a q_b q_c (p_{-a} + p_c)$. By Axiom C2, we can divide with q_b, giving $q_a q_c (p_a + p_{-c}) \succsim' q_a q_b (p_{-a} + p_c)$, ie, we have shown $a \succsim c$. This proves transitivity, ie, $\succsim$ is a complete ordering on $\mathcal{C}$.

To prove Axiom C1 on $\mathcal{C}_\mathbb{Q}$, note that the identity $a = p_a - p_{-a}$ gives $a - b = p_{a-b} - p_{b-a}$ and $a - b = p_a - p_{-a} - p_b + p_{-b}$, resulting in the identity

$$p_{a-b} + p_{-a} + p_b = p_{b-a} + p_a + p_{-b}.$$

Together with Axiom C1 and C2 and the definition of $\succsim$, this gives

$$\begin{aligned}
a \succsim b \quad &\Leftrightarrow \quad q_a q_b (p_a + p_{-b}) \succsim' q_a q_b (p_{-a} + p_b) \\
&\Leftrightarrow \quad q_a q_b (p_{a-b} + p_a + p_{-b}) \succsim' q_a q_b (p_{a-b} + p_{-a} + p_b) \\
&\Leftrightarrow \quad q_a q_b (p_{a-b} + p_a + p_{-b}) \succsim' q_a q_b (p_{b-a} + p_a + p_{-b}) \\
&\Leftrightarrow \quad q_a q_b p_{a-b} \succsim' q_a q_b p_{b-a} \\
&\Leftrightarrow \quad q_{a-b} p_{a-b} \succsim' q_{a-b} p_{b-a} \\
&\Leftrightarrow \quad a - b \succsim 0.
\end{aligned}$$

From this equivalence, we immediately get $a \succsim b$ iff $a + c \succsim b + c$, ie, $\mathcal{C}_{\mathbb{Q}}$ satisfies Axiom C1.

To prove Axiom C2, it suffices to show that for any $a \in \mathcal{C}_{\mathbb{Q}}$ and positive integer k, we have $a \succsim 0$ iff $ka \succsim 0$. We can prove this by combining Axiom C2 on $\mathcal{C}$ with the identity $kp_a = p_{ka}$ for any $k > 0$:

$$\begin{aligned}
a \succsim 0 \quad &\Leftrightarrow \quad q_a p_a \succsim' q_a p_{-a} \quad &&\Leftrightarrow \quad k q_a p_a \succsim' k q_a p_{-a} \\
&\Leftrightarrow \quad q_a p_{ka} \succsim' q_a p_{-ka} \quad &&\Leftrightarrow \quad ka \succsim kb
\end{aligned}$$

To prove Axiom C3, it suffices to show that $a \succsim 0$ and $b \succsim 0$ implies $a + b \succsim 0$. So assume $a \succsim 0$ and $b \succsim 0$, then $q_a p_a \succsim' q_a p_{-a}$ and $q_b p_b \succsim' q_b p_{-b}$, so Axiom C2 and C3 give

$$q_a q_b (p_a + p_b) \succsim' q_a q_b (p_{-a} + p_{-b})$$

The identity $p_{a+b} + p_{-a} + p_{-b} = p_{-a-b} + p_a + p_b$ combined with Axiom C1 and C2 then gives

$$\begin{aligned}
&q_a q_b (p_{a+b} + p_a + p_b) \succsim' q_a q_b (p_{a+b} + p_{-a} + p_{-b}) \\
\Rightarrow \quad &q_a q_b (p_{a+b} + p_a + p_b) \succsim' q_a q_b (p_{-a-b} + p_a + p_b) \\
\Rightarrow \quad &q_a q_b p_{a+b} \succsim' q_a q_b p_{-a-b} \\
\Rightarrow \quad &q_{a+b} p_{a+b} \succsim' q_{a+b} p_{-a-b} \\
\Rightarrow \quad &a + b \succsim 0
\end{aligned}$$

To prove Axiom C4, it suffices to show that $a \succ 0$ and $b \in \mathcal{C}_{\mathbb{Q}}$ implies that there exists a positive integer k such that $ka \succ b$. Suppose this does not hold, then there exists $a, b \in \mathcal{C}_{\mathbb{Q}}$ such that $a \succ 0$ and $ka \prec b$ for all positive integers k; since $p_b \succsim b$, we may assume $p_{-b} = 0$ without any loss of generality. We then have $q_a p_a \succ q_a p_{-a}$ and hence $q_a q_b p_a \succ q_a q_b p_{-a}$, but $q_{ka} q_b p_{ka} \prec q_{ka} q_b (p_{-ka} + p_b)$ and hence $k \cdot q_a q_b p_a \prec k \cdot q_a q_b p_{-a} + q_a q_b p_b$ for all $k \in \mathbb{N}$, contradicting Axiom C4 in $\mathcal{C}$.

We have proved that the ordering on $\mathcal{C}_{\mathbb{Q}}$ is complete, satisfies Axioms C1–C4, and coincides with the ordering on $\mathcal{C}$, ie, the proposition holds. □

Since we can replace a with $-a$ if $a \prec 0$, we will assume without loss of generality that $a \succsim 0$ for all $a \in \mathcal{A}$. With this assumption, we will now extend $\mathcal{C}_\mathbb{Q}$ to the corpus set $\mathcal{C}_\mathbb{R}$.

Definition 8 Let $\mathcal{C}_\mathbb{R}$ be the corpus set induced by $\mathcal{A}$ over $\mathbb{R}$, with $a \succsim' 0$ for all $a \in \mathcal{A}$ where $\succsim'$ denotes the ordering on $\mathcal{C}_\mathbb{Q}$. Given $c \in \mathcal{C}_\mathbb{R}$, let

$$L_n(c) = \frac{1}{n} \sum_{a \in \mathcal{A}} \lfloor nc(a) \rfloor a$$

where $\lfloor r \rfloor$ denotes the largest integer below r.[1] Define an ordering $\succsim$ on $\mathcal{C}_\mathbb{R}$ by defining $c \succsim c'$ iff $c - c' \succsim 0$, and

$$c \succsim 0 \qquad \text{iff} \qquad \forall m\, \exists n > m \colon L_n(c) \succsim' -\frac{u_c}{m}$$

where u_c is the corpus defined by $u_c(a) = 1$ if $c(a) \neq 0$, and $u_c(a) = 0$ if $c(a) = 0$. Equivalently, $c \prec 0$ iff $\exists m\, \forall n > m \colon L_n(c) \prec' -\frac{u_c}{m}$.

Rather than giving a direct proof that $\mathcal{C}_\mathbb{R}$ is a corpus space, we prove that $\mathcal{C}_\mathbb{R}$ is a mixture set that satisfies Herstein and Milnor's axioms of utility from Theorem 1. We start by proving the following lemma.

Lemma 3 *Let $c \in \mathcal{C}_\mathbb{R}$, and let m, n be positive integers. Then*

$$-\frac{u_c}{\min(m,n)} \prec' L_m(c) - L_n(c) \prec' \frac{u_c}{\min(m,n)}.$$

Proof. Since $x - 1 < \lfloor x \rfloor \leq x$, we have

$$\frac{1}{m}\lfloor ms \rfloor - \frac{1}{n}\lfloor ns \rfloor < \frac{ms}{m} - \frac{ns-1}{n} = \frac{1}{n} \leq \frac{1}{\min(m,n)}.$$

Since $a \succsim' 0$ for all $a \in \mathcal{A}$, we therefore have

$$L_m(c) - L_n(c) = \sum_{a \in \mathcal{A}} \left(\frac{1}{m}\lfloor mc(a) \rfloor - \frac{1}{n}\lfloor nc(a) \rfloor \right) a \prec' \frac{u_c}{\min(m,n)}$$

The lemma follows immediately by exchanging m and n. □

Proposition 4 *$\mathcal{C}_\mathbb{R}$ is a completely ordered mixture space.*

Proof. It follows immediately from the definition of $\mathcal{C}_\mathbb{R}$ that it is a mixture space. To prove that $\succsim$ is well-ordered, it suffices to prove that $c \succsim 0$ or $c \precsim 0$ for all $c \in \mathcal{C}_\mathbb{R}$. Suppose we can find $c \in \mathcal{C}_\mathbb{R}$ such that $c \succ 0$ and $c \prec 0$, then $\exists m_1\, \forall n > m_1 \colon L_n(c) \succ' u_c/m_1$ and $\exists m_2\, \forall n > m_2 \colon L_n(c) \prec' -u_c/m_2$, so for $n > \max(m_1, m_2)$ we have

[1] Our definition of $L_n(c)$ only works under the assumption that $a \succsim 0$ for all $a \in \mathcal{A}$. To deal with general $\mathcal{A}$, it is necessary to replace our definition of $L_n(c)$ with

$$L_n(c) = \frac{1}{n} \sum_{a \in \mathcal{A}} \ell_a(nc(a))a$$

where $\ell_a(r) = \lfloor r \rfloor$ if $a \succsim 0$, and $\ell_a(r) = \lceil r \rceil = -\lfloor -r \rfloor$ if $a \prec 0$. Similarly, we must define $u_c(a) = 1$ if $a \succsim 0$, and $u_c(a) = -1$ if $a \prec 0$.

$L_n(c) \succ' u_c/m_1 \succ' 0$ and $L_n(c) \prec' -u_c/m_2 \prec' 0$, a contradiction since $\succsim'$ is well-ordered.

To prove that $\succsim$ is transitive, it suffices to prove that $c \succsim 0$ and $c' \succsim 0$ implies $c + c' \succsim 0$. So suppose $c \succsim 0$ and $c' \succsim 0$, and let m be given. Then we can find $n > 2m$ and $n' > 2m$ such that $L_n(c) \succsim -u_c/2m$ and $L_{n'}(c') \succsim -u_{c'}/2m$. By Lemma 3, $L_n(c') - L_{n'}(c') \succ' -u_{c'}/2m$. Adding the three inequalities together, we therefore get:

$$L_n(c) + L_{n'}(c') \succ' -\frac{u_c}{2m} - \frac{u_{c'}}{2m} - \frac{u_{c'}}{2m} \succsim' -\frac{u_c + u_{c'}}{m}$$

Since $u_c + u_{c'} \succsim u_{c+c'}$, and $L_n(c + c') \succsim L_n(c) + L_n(c')$ follows from $\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor$, the inequality above shows that

$$L_n(c + c') \succsim' -\frac{u_{c+c'}}{m}.$$

That is, we have proven that $c + c' \succsim 0$. □

Proposition 5 *$\mathcal{C}_\mathbb{R}$ is an extension of $\mathcal{C}_\mathbb{Q}$, ie, the ordering $\succsim$ on $\mathcal{C}_\mathbb{R}$ coincides with the ordering $\succsim'$ on $\mathcal{C}_\mathbb{Q}$ when restricted to $\mathcal{C}_\mathbb{Q}$.*

Proof. It suffices to prove that if $c \in \mathcal{C}_\mathbb{Q}$, then $c \succsim' 0$ implies $c \succsim 0$. So suppose $c \succsim' 0$. Then $q_c c(a)$ is an integer for all $a \in \mathcal{A}$, so $L_n(c) = c$ if n is a multiple of q_c. It follows immediately that for all m, there exists a multiple n of q_c such that $n > m$ and $L_n(c) = c \succsim 0 \succsim -u_c/k$, ie, we have proven that $c \succsim 0$. □

The substitutability axiom follows almost immediately from the following lemma.

Lemma 6 *If $c \succsim 0$ in $\mathcal{C}_\mathbb{R}$ and $r > 0$, then $rc \succsim 0$.*

Proof. Suppose $c \succsim 0$ and $r \in]0, 1[$, and let m be given. Choose $n > 2m$ such that $L_n(c) \succsim -u_c/2m$. Setting $x = i + s = i'/r + s'$ where i, i' are integers and $s \in [0, 1[$, $s' \in [0, 1/r[$, we see that $\lfloor rx \rfloor - r\lfloor x \rfloor = i' - ri = r(s - s') \geq -1$. Using $\lfloor x \rfloor \leq x$, we calculate

$$\frac{\lfloor nrc \rfloor}{n} - \frac{\lfloor nr \rfloor}{n} \cdot \frac{\lfloor nc \rfloor}{n} \geq \frac{\lfloor nrc \rfloor - r\lfloor nc \rfloor}{n} \geq -\frac{1}{n} \geq -\frac{1}{2m}.$$

We therefore have

$$L_n(rc) - \frac{\lfloor nr \rfloor}{n} \cdot L_n(c) \succsim -\frac{u_c}{2m}$$

Multiplying the inequality $L_n(c) \succsim -u_c/2m$ with $\lfloor nr \rfloor/n$, and using $\lfloor nr \rfloor/n < 1$ when $r \in]0, 1[$, we get

$$\frac{\lfloor nr \rfloor}{n} \cdot L_n(c) \succsim -\frac{u_c}{2m} \cdot \frac{\lfloor nr \rfloor}{n} \succsim -\frac{u_c}{2m}.$$

Adding the two inequalities, we get

$$L_n(rc) \gtrsim -\frac{u_c}{m},$$

which proves that $rc \gtrsim 0$. The proof of Lemma 4 shows that for all positive integers k, $c \gtrsim 0$ implies $kc = c + \cdots + c \gtrsim 0$. Writing $r = k + r'$ where k is an integer and $r \in [0, 1[$, $c \gtrsim 0$ implies $kc \gtrsim 0$ and $r'c \gtrsim 0$, and hence $kc + r'c = rc \gtrsim 0$. This proves the lemma. □

Proposition 7 *$\mathcal{C}_{\mathbb{R}}$ satisfies Herstein and Milnor's weak substitutability axiom, ie, if $a \sim a'$, then $\frac{1}{2}a + \frac{1}{2}b \sim \frac{1}{2}a' + \frac{1}{2}b$ for all $a, a', b \in \mathcal{C}_{\mathbb{R}}$.*

Proof. To prove the proposition, we note that $a - a' \gtrsim 0$ implies $\frac{1}{2}(a - a') = \frac{1}{2}a + \frac{1}{2}b - (\frac{1}{2}a' + \frac{1}{2}b) \gtrsim 0$ by Lemma 6. Exchanging $\gtrsim$ with $\lesssim$, the proposition follows immediately. □

Herstein and Milnor's continuity axiom follows almost immediately from the following lemma.

Lemma 8 *$\mathcal{C}_{\mathbb{R}}$ satisfies Axiom C4, ie, if $a \succ 0$ and $b \in \mathcal{C}_{\mathbb{R}}$, then there exists a positive integer k such that $ka \succ b$.*

Proof. Suppose $a \succ 0$, then there exists m such that $a' = L_m(a) \succ 0$, and since $L_m(a) \lesssim a$, we have $0 \prec a' \lesssim a$. Similarly, $b' = -L_1(-b)$ satisfies $b' \gtrsim b$. Since $a', b' \in \mathcal{C}_{\mathbb{Q}}$ and $a' \succ 0$, it follows from Axiom C4 on $\mathcal{C}_{\mathbb{Q}}$ that there exists k such that $ka' \succ b'$. Since $a \gtrsim a'$, we have $ka \gtrsim ka'$, and combining this with $b' \gtrsim b$, transitivity therefore gives:

$$ka \gtrsim ka' \succ b' \gtrsim b.$$

□

Proposition 9 *$\mathcal{C}_{\mathbb{R}}$ satisfies Herstein and Milnor's continuity axiom, ie, the sets $\{\lambda \mid \lambda a + (1-\lambda)b \gtrsim c\}$ and $\{\lambda \mid \lambda a + (1-\lambda)b \lesssim c\}$ are closed for all $a, b, c \in \mathcal{C}_{\mathbb{R}}$.*

Proof. Define $U_{a,b,c} = \{\lambda \in \mathbb{R} \mid \lambda a + (1-\lambda)b \succ c\}$. Since closed sets are defined as complements of open sets, Herstein and Milnor's continuity axiom follows if we can prove that $U_{a,b,c}$ is open for all $a, b, c \in \mathcal{C}_{\mathbb{R}}$. Suppose this is not the case, then we can find $a, b, c \in \mathcal{C}_{\mathbb{R}}$ such that $U_{a,b,c}$ is not open, ie, there exists $\lambda \in U_{a,b,c}$ such that for all $\epsilon > 0$ we can find $\delta \in]-\epsilon, \epsilon[$ such that $\lambda + \delta \notin U_{a,b,c}$. Write $s = \lambda a + (1-\lambda)b - c$ and $s' = a - b$. Then $\lambda \in U_{a,b,c}$ implies $s \succ 0$, and $\lambda + \delta \notin U_{a,b,c}$ implies $s + \delta s' \lesssim 0$, so obviously $\delta s' \prec 0$. We may assume without loss of generality that $s' \succ 0$. By Lemma 6, we then have $\delta s' \succ -\epsilon s'$, and hence $0 \gtrsim s + \delta s' \succ s - \epsilon s'$. In particular, setting $\epsilon = \frac{1}{k}$ where k is a positive integer, we have $s - \frac{1}{k}s' \prec 0$, so by Lemma 6, we have $ks - s' \prec 0$ for all positive integers k. This contradicts that $s \succ 0$ and

Axiom C4 implies that $ks \succ s'$ for some k. From this contradiction, we conclude that $U_{a,b,c}$ is open for all $a, b, c \in \mathcal{C}_{\mathbb{R}}$, ie, Herstein and Milnor's continuity axiom holds. □

The reader is referred to Munkres (1975) for the definition of open and closed sets, continuity, etc. Propositions 4, 7, and 9 demonstrate that $\mathcal{C}_{\mathbb{R}}$ satisfies the axioms of utility, so from Theorem 1, we can conclude that the ordering on $\mathcal{C}_{\mathbb{R}}$ can be expressed by means of a measurable utility. From the existence of a measurable utility, it immediately follows that $\mathcal{C}_{\mathbb{R}}$ satisfies Axioms C1–C4, ie, that $\mathcal{C}_{\mathbb{R}}$ is a corpus space. We therefore have:

Theorem 10 *$\mathcal{C}_{\mathbb{R}}$ is a corpus space, and the ordering on $\mathcal{C}_{\mathbb{R}}$ can be expressed by means of a measurable utility, ie, there exists a linear order-preserving mapping $u \colon \mathcal{C}_{\mathbb{R}} \to \mathbb{R}$.*

From Propositions 2 and 5 and Theorem 10, we have:

Theorem 11 *Let $\mathcal{C}$ be a corpus space over $\mathbb{N}_0$ with analyses $\mathcal{A}$. Then $\mathcal{C}$ can be extended to corpus spaces $\mathcal{C}_{\mathbb{Q}}$ and $\mathcal{C}_{\mathbb{R}}$, and the ordering on $\mathcal{C}$, $\mathcal{C}_{\mathbb{Q}}$, and $\mathcal{C}_{\mathbb{R}}$ can be expressed by means of a measurable utility $u \colon \mathcal{C}_{\mathbb{R}} \to \mathbb{R}$.*

We have thereby proven that if the reader accepts our axiomatic model of speaker preferences, then the reader is also forced to accept that these preferences can be encoded by means of a measurable utility.

1.4 Discussion and concluding remarks

The axiomatic model of speaker preferences suggests that instead of viewing a grammar as a device that generates all the grammatical sequences in a language, and none of the ungrammatical ones, as proposed by Chomsky (1957, p. 13), it is better to view a grammar as a device that computes the measurable utility that encodes a speaker's linguistic preferences. Language acquisition can then be conceived as the problem of reconstructing a measurable utility that encodes the observed preferences of other speakers, and language production and language understanding can be conceived as optimization problems where the goal is to find an optimal analysis with a given meaning or a given speech signal.

The chomskyan binary distinction between grammatical and ungrammatical sentences can in principle be encoded as a measurable utility that maps every analysis to either 0 or 1. However, this constitutes a very poor model of speaker preferences because it leads to the prediction that all grammatical analyses are equally good, and all ungrammatical analyses are equally bad. Human speaker preferences are obviously far more fine-grained than that. For example, in disambiguation tasks, human speakers can select the most plausible analysis among the many possible grammatical and ungrammatical analyses, and human speakers are also capable of analyzing ungrammatical input — a task that would be impossible if all ungrammatical analyses were equally bad.

These insights are not new. Indeed, the log-probabilities in probabilistic grammars such as (Collins, 1997) can be viewed as special instances of measurable utilities. However, in many linguistic theories, including GB, LFG, and HPSG, probabilities or other kinds of measurable utilities are not viewed as a core aspect of the theory — if they are used at all, they are at best viewed as practical add-ons used by computational linguists for disambiguation, and little attention is paid to ensuring that the probabilistic add-ons are good models of human speaker preferences on their own. Our proof shows that if our axioms provide a good model of speaker preferences, as we have argued that they do, then all grammars can be expressed as measurable utilities. That is, our results suggest that measurable utilities deserve a far more prominent place in linguistic theory than today as the fundamental expression of a speaker's preferences.

References

Chomsky, Noam. 1957. *Syntactic structures*. Mouton, The Hague.

Collins, Michael. 1997. Three generative, lexicalized models for statistical parsing. In P. R. Cohen and W. Wahlster, eds., *Proceedings of the Thirty-Fifth Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, pages 16–23. Somerset, New Jersey: Association for Computational Linguistics.

Herstein, I. N. and John Milnor. 1953. An axiomatic approach to measurable utility. *Econometrica, Journal of the Econometric Society* 21(2).

Munkres, James R. 1975. *Topology. A first course*. Prentice Hall.

Ramsey, Frank P. 1931. Truth and probability. In R. B. Braithwaite, ed., *The foundations of mathematics and other logical essays*. Harcourt Brace Jovanovich.

Russell, Stuart J. and Peter Norvig. 1995. *Artificial intelligence*. Prentice-Hall.

Sorace, Antonella and Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115(11):1497–1524.

von Neumann, John and Oskar Morgenstern. 1944. *Theory of games and economic behavior*. Princeton University Press.